



SEP  
SECRETARÍA DE  
EDUCACIÓN PÚBLICA



---

TECNOLÓGICO NACIONAL DE MÉXICO  
INSTITUTO TECNOLÓGICO DE APIZACO

---

**“CMMI DEV 2 COMO BASE DEL PROCESO  
ETL PARA LA GENERACIÓN DE  
CONOCIMIENTO EN LA INTELIGENCIA DE  
NEGOCIOS”**

---

T E S I S

QUE PARA OBTENER EL GRADO DE:  
MAESTRA EN SISTEMAS COMPUTACIONALES

PRESENTA:

ING. RAQUEL CONTRERAS DE LA FUENTE

DIRECTOR: DR. EDMUNDO BONILLA HUERTA

AGOSTO 2018, APIZACO TLAXCALA

## Comité Tutorial

### **Codirector**

Profesor Dr. Perfecto Malaquías Quintero Flores

### **Tutor**

Profesor Dr. José Federico Ramírez Cruz

### **Revisor**

Profesor Dr. José Crispin Hernández Hernández

Apizaco, Tlax., 25 de Junio de 2018

No. de Oficio: DEPI/248/18

ASUNTO: Se Autoriza Impresión de Tesis de Grado.

**ING. RAQUEL CONTRERAS DE LA FUENTE**  
CANDIDATA AL GRADO DE MAESTRA  
EN SISTEMAS COMPUTACIONALES  
No. de Control: **M11370777**  
PRESENTE.

Por este medio me permito informar a usted, que por aprobación de la Comisión Revisora asignada para valorar el trabajo, mediante la Opción: **I Tesis de Grado por Proyecto de Investigación**, de la **Maestría en Sistemas Computacionales**, que presenta con el tema: **"CMMI DEV 2 COMO BASE DEL PROCESO ETL PARA LA GENERACIÓN DE CONOCIMIENTO EN LA INTELIGENCIA DE NEGOCIOS"** y conforme a lo establecido en el Procedimiento para la Obtención del Grado de Maestría en el Instituto Tecnológico, la División de Estudios de Posgrado e Investigación a mi cargo le emite la:

**AUTORIZACIÓN DE IMPRESIÓN**

Debiendo entregar un ejemplar del mismo debidamente encuadernado y seis copias en CD en formato PDF, para presentar su Acto de Recepción Profesional a la brevedad.

Sin otro particular por el momento, le envío un cordial saludo.

**ATENTAMENTE**  
EXCELENCIA EN EDUCACIÓN TECNOLÓGICA®  
PENSAR PARA SERVIR, SERVIR PARA TRIUNFAR®

**DR. JOSÉ FEDERICO CASCO VÁSQUEZ**  
JEFE DE LA DIVISIÓN DE ESTUDIOS DE  
POSGRADO E INVESTIGACIÓN.



SECRETARÍA DE EDUCACIÓN PÚBLICA  
TECNOLÓGICO NACIONAL  
DE MÉXICO  
INSTITUTO TECNOLÓGICO DE APIZACO  
DIVISIÓN DE ESTUDIO  
DE POSGRADO E INVESTIGACIÓN

JFCV/MSH\*mebr.

C.p. Expediente.

Apizaco, Tlax., 25 de Junio de 2018

ASUNTO: Aprobación del trabajo de Tesis de Maestría.

DR. JOSÉ FEDERICO CASCO VÁSQUEZ  
JEFE DE LA DIVISIÓN DE ESTUDIOS DE  
POSGRADO E INVESTIGACIÓN.  
P R E S E N T E.

Por este medio se le informa a usted, que los integrantes de la **Comisión Revisora** para el trabajo de tesis de maestría que presenta la **ING. RAQUEL CONTRERAS DE LA FUENTE**, con número de control **M11370777**, candidata al grado de **Maestra en Sistemas Computacionales** y egresada del **Instituto Tecnológico de Apizaco**, cuyo tema es "**CMMI DEV 2 COMO BASE DEL PROCESO ETL PARA LA GENERACIÓN DE CONOCIMIENTO EN LA INTELIGENCIA DE NEGOCIOS**", fue:

**A P R O B A D O**

Lo anterior, al valorar el trabajo profesional presentado por la candidata y constatar que las observaciones que con anterioridad se le marcaron así como correcciones sugeridas para su mejora ya han sido realizadas.

Por lo que se avala se continúe con los trámites pertinentes para su titulación.

Sin otro particular por el momento, le envió un cordial saludo.

LA COMISIÓN REVISORA

  
DR. EDMUNDO BONILLA HUERTA

  
DR. PERFECTO MALAQUIAS QUINTERO FLORES

  
DR. JOSÉ FEDERICO RAMÍREZ CRUZ

  
DR. JOSÉ CRISPIN HERNÁNDEZ HERNÁNDEZ

C. p.- Interesada

# Agradecimientos

"El mejor maestro es aquel que te dice dónde buscar, pero no te dice qué ver."

Alexandra K. Trenfor

Me gustaría agradecer especialmente a mi director de tesis el Doctor **Edmundo Bonilla Huerta** por el gran apoyo que siempre me brindo, por su paciencia, confianza y por brindarme la oportunidad de vivir esta grata experiencia.

A mis padres **Rosario y Luis Antonio** por ser las personas que me impulsan a continuar en el camino del aprendizaje, sin duda alguna ustedes son el gran apoyo que Dios me dio y que necesitó para salir adelante, mil gracias a mi madre por creer en mi por apoyarme aún cuando las circunstancias eran difíciles, sin duda tú has sido la persona que más estímulo me dio, gracias a ti aprendí a llegar lejos.

A la Maestra **Lorena Roldan** por su amabilidad al ofrecerme su apoyo incondicional para la revisión de tesis, gracias por tomarse el tiempo de colaborar conmigo.

Sin olvidar a **Conacyt** ya que esto no hubiera sido posible sin el apoyo que representa, gracias a esta organización que contribuye al crecimiento profesional de tantos jóvenes que desean superarse continuamente.

# Resumen

El proyecto que se presenta a continuación está enfocado en su totalidad a la colaboración con el crecimiento de la empresa Miracle Business Network dentro del sector de tecnologías de la información, dio inicio en el año 2000 en la ciudad de México, en el comienzo MBN brindó servicios de administración de base de datos a Oracle de México con ello se colocó entre las redes de partners de Oracle, en el año 2006 inicia operaciones en el estado de Tlaxcala, entre las empresas a las que le brindó servicios están: Pemex, Televisa y Nextel. En 2008 logra la implantación de la norma MoProSoft y en 2015 Software Engineering Institute (SEI) otorga el nivel de madurez 2 del modelo de CMMI-DEV v1.3. En 2016 es galardonada con el premio tlaxcalteca a la competitividad, por lo anterior la empresa es una de las organizaciones con mayor crecimiento en el estado.

MBN se mantiene a la vanguardia y debido a ello a buscado la implantación de nuevas tecnologías como lo es la inteligencia de negocios dentro de la organización, que hasta ahora ha evolucionado constantemente convirtiéndose en un campo de aplicación muy grande e incierto, pero con ello surgen las nuevas aplicaciones con herramientas que permiten la implantación exitosa de esta tecnología sin embargo el costo de dicha implementación es algo a lo que no todas las organizaciones pueden afrontar y se desconoce si alguna compañía del sector de TI haya sido capaz de implementar esta tecnología en la misma organización.

El trabajo que representa el proceso ETL dentro de la inteligencia de negocios

---

es bastante complejo por lo que se requiere del conocimiento total en la manera de operar en la organización, por lo tanto, el trabajo principal es la investigación y a continuación la aplicación ya que el panorama general en la organización carece de antecedentes concretos que puedan llevar a la solución óptima del procesamiento de la información que genera la empresa.

Durante el desarrollo de esta tesis se exploraron diferentes herramientas, entre ellas Talend Open Studio para decretar la que pueda adaptarse más a las necesidades de la empresa y a su vez se exploró el software R como una herramienta externa que pueden significar la reducción en los gastos en licencias de las aplicaciones ya existentes. También se propone un enfoque diferente en base a Indicadores clave de rendimiento (KPIS) que puedan favorecer la limpieza, la extracción y la transformación de la información para dar una solución más precisa y consistente de aquellas actividades cruciales para considerar durante la toma de decisiones.

# Abstract

The following project presented is focused in its entirety on collaboration with the growth of the company Miracle business network, that is positioned in the information technology sector, MBN started in the year 2000 in Mexico City, in the beginning MBN provided data bases administration services to Oracle of Mexico with that it established itself among network partners at Oracle, in 2006 it started operation in the state of Tlaxcala, among the Enterprises to which it served were the following: Pemex, Televisa, and Nextel. In 2008 it achieved the implemented the Microsoft certification standard and in 2015 the Software Engineering Institute (SEI) grants level 2 in CMMI-DEV v1.3. model. In 2016 it is awarded with Tlaxcalteca prize for competitiveness because of this the company is one of the Enterprises with more growth in the state.

mbn is at the forefront and due to this it has searched to implement new technologies like business intelligence in organization, which until now has constantly evolved into a very large and uncertain field of application, but with that new applications with tools that allow the successful implementation of this technology arise however the cost of such implementation is something that not all organizations can face and it is unknown if any company in the IT sector has been able to implement this technology in the same organization.

The work that the ETL process represents within the business intelligence is quite complex so it requires the total knowledge in the way the it is operated in the

---

organization, therefore, the main job is the investigation and then the application since the The general panorama in the organization lacks a specific background that can lead to the optimal solution for processing the information generated by the company.

During the development of this thesis, different tools were explored, among them Talend Open Studio to decree the one that can be adapted better to the needs of the company and at the same time the R software was explored as an external tool that can mean the reduction in expenses in licenses of existing applications. A different approach is also proposed based on Key Performance Indicators (KPIS) that can favor the cleaning, extraction and transformation of information to give a more precise and consistent solution of those crucial activities to consider during decision making.

# Índice general

<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	Descripción del problema . . . . .	1
1.2	Justificación . . . . .	2
1.3	Objetivo general . . . . .	3
1.4	Objetivos específicos . . . . .	3
1.5	Pregunta de investigación . . . . .	3
1.6	Descripción del documento . . . . .	4
1.7	Conclusión capítulo 1 . . . . .	4
<b>2</b>	<b>Marco teórico</b>	<b>5</b>
2.1	Conceptos . . . . .	5
2.2	Herramientas ETL . . . . .	8
2.2.1	Talend open studio para la integración de datos . . . . .	8
2.2.2	ETL Informatica powercenter . . . . .	8
2.2.3	IBM Datastage . . . . .	9
2.2.4	Oracle data Integrator . . . . .	10
2.3	Tabla comparativa de herramientas para el proceso ETL . . . . .	11
2.4	R Software libre . . . . .	12
2.5	Estado del arte: proceso ETL y key performance indicators . . . . .	13
2.6	Conclusión capítulo 2 . . . . .	17

<b>3 Metodología</b>	<b>19</b>
3.1 Modelo de madurez de capacidad integrado (CMMI) . . . . .	19
3.1.1 Componentes del modelo . . . . .	21
3.1.2 Nivel de madurez 2 Gestionado . . . . .	22
3.1.3 Categoría de las áreas de proceso . . . . .	23
3.1.4 Objetivos de las Áreas de proceso Nivel 2 . . . . .	23
3.2 Key performance Indicators . . . . .	30
3.3 Pasos para crear KPIS . . . . .	31
3.4 Conclusión capítulo 3 . . . . .	35
<b>4 Extracción Transformación y carga</b>	<b>37</b>
4.1 ETL en Business Intelligence . . . . .	37
4.2 Pasos para ETL . . . . .	39
4.3 Metodología ETL para la organización . . . . .	49
4.4 Conclusión capítulo 4 . . . . .	51
<b>5 Resultados de KPIS y Pruebas ETL</b>	<b>53</b>
5.1 Resultado de KPI definidas en la organización según CMMI . . . . .	53
5.2 Áreas de proceso que proveen información para la solución de KPIS .	64
5.3 Extracción de KPIS con R . . . . .	65
5.3.1 Extracción de KPI 1 . . . . .	66
5.3.2 Extracción de KPI 2 . . . . .	67
5.3.3 Extracción de KPI 3 . . . . .	68
5.3.4 Extracción de KPI 4 . . . . .	69
5.3.5 Extracción de KPI 5 . . . . .	70
5.3.6 Extracción de KPI 6 . . . . .	71
5.3.7 Extracción de KPI 9 . . . . .	72
5.3.8 Extracción de KPI 10 . . . . .	73
5.4 Conclusión capítulo 4 . . . . .	74

<b>6 Conclusiones y trabajos futuros</b>	<b>75</b>
6.1 Conclusiones . . . . .	75
6.2 Trabajos Futuros . . . . .	76
<b>7 Apendice</b>	<b>79</b>
<b>8 Bibliografia</b>	<b>89</b>

# Índice de figuras

3.1. Constelaciones de CMMI. . . . .	20
3.2. Áreas de Proceso CMMI DEV. . . . .	21
3.3. Áreas de proceso del nivel 2. . . . .	23
3.4. Clasificación de las áreas de proceso del nivel 2 en base a su categoría. . . . .	23
3.5. Documentación para estimar el alcance del proyecto. . . . .	28
3.6. Documentación para el seguimiento de cambios. . . . .	29
3.7. Documentación para la evolución del producto. . . . .	29
3.8. Pasos útiles para la generación de KPIS. . . . .	34
4.1. Fases del proceso BI. . . . .	38
4.2. Datos importantes para el proceso de extracción. . . . .	39
4.3. Actividades de la transformación. . . . .	40
4.4. Actividades a considerar para la carga. . . . .	42
4.5. Interacción de los pasos para el proceso ETL. . . . .	42
4.6. Los cuatro grupos del proceso ETL. . . . .	43
4.7. Proceso ETL para MBN . . . . .	50
5.1. Diagrama KPI ¿Qué tan efectivo es un proyecto? . . . . .	65
5.2. Documentación original MBN . . . . .	65
5.3. Datos Extraídos del cronograma de actividades. . . . .	66
5.4. Datos extraídos del reporte de actividades. . . . .	66
5.5. Documentación del plan de proyecto. . . . .	67
5.6. Datos extraídos del plan de proyecto. . . . .	67
5.7. Documentación del plan general y datos obtenidos. . . . .	68

5.8. Documentación matriz de trazabilidad. . . . .	69
5.9. Datos obtenidos de la matriz de trazabilidad. . . . .	69
5.10. Documentación plan de pruebas. . . . .	70
5.11. Datos obtenidos del plan de pruebas. . . . .	70
5.12. Documentación seguimiento de hitos. . . . .	71
5.13. Procesamiento wbs e hitos. . . . .	71
5.14. Documentación carta de aceptación. . . . .	72
5.15. Datos obtenidos de la carta de aceptación. . . . .	72
5.16. Documentacion Reporte de actividades. . . . .	73
5.17. Datos obtenidos de seguimiento de actividades. . . . .	73

# Índice de tablas

2.1. Comparación de herramientas . . . . .	11
5.1. Estado actual del proyecto. . . . .	54
5.2. Quejas generadas durante el proyecto. . . . .	55
5.3. Habilidades para conformación del equipo. . . . .	56
5.4. Uso del proyecto. . . . .	57
5.5. Estimación de Tecnología. . . . .	58
5.6. Calidad del proyecto. . . . .	59
5.7. Aceptación del usuario. . . . .	60
5.8. Efectividad en la ejecución. . . . .	61
5.9. Efectividad en la ejecución. . . . .	62
5.10. Tasa de cumplimiento de actividades. . . . .	63
5.11. Procesos que proveen información. . . . .	64

# 1 Introducción

El nivel de transformación de las empresas especialmente del sector de Tecnologías de la Información (TI) ha ido en aumento en los últimos años, por lo que se han visto en la necesidad de mejorar los servicios que ofrecen a sus clientes y a la vez garanticen la confiabilidad de los mismos. Debido a ello las organizaciones tienden a la adopción de normas de calidad. Miracle Bussines Network, empresa líder en el estado de Tlaxcala dedicada a la creación de software a la medida y al soporte de tecnologías de información, comprometida con el servicio al cliente cuenta con certificaciones enfocadas en garantizar la mejora en la calidad de las etapas de procesos de desarrollo para proyectos. La manipulación de este tipo de certificaciones genera información relevante que muchas veces no es explotada de forma adecuada para permitir una buena práctica en la toma de decisiones, MBN actualmente cuenta con el nivel 2 en la escala de certificaciones de CMMI para el desarrollo. El diseño de la herramienta de Business Intelligence (BI), tiene como base el proceso ETL para lograr la asertividad y confiabilidad en la información recabada por este proceso.

## 1.1. Descripción del problema

El incremento de la información ha sido relevante para la mejora de procesos dentro de las organizaciones ya que se generan datos importantes que requieren de análisis para obtener beneficios durante el desarrollo de proyectos, esta información es de tipo semiestructurada y no estructurada, por lo cual se requiere del conocimiento eficiente que permita una toma de decisiones que aporte datos adecuados y asertivos para ejercer una mejora. MBN requiere de la información relevante que generan los

procesos de fábrica de software y Administración de proyectos Específicos (APE), ambos bajo la norma MoProsoft, sin embargo, la información de estos procesos se produce mediante la certificación de calidad CMMI y las áreas de proceso enfocadas al desarrollo en el nivel 2, que son: Gestión de requerimientos (REQM), Planificación del proyecto (PP), Monitoreo y control del proyecto (PMC), Gestión de configuración (CM), Aseguramiento de la calidad de proceso y de producto (PPQA) y Medición y análisis (MA). Estos procesos almacenan información por medio de una base de conocimientos que está conformada mediante diferentes tipos de archivos, en su mayoría docx, pptx, xlsx y pdf., es debido a ello que existe una cantidad relevante de información, pero es poco efectiva por lo cual conlleva a la falta de asertividad y el retraso en la generación de informes que contribuyan de manera eficiente a la obtención de conocimiento.

### **1.2. Justificación**

El primer paso y el más complejo para la implementación de Business Intelligence es el proceso de extracción, transformación y carga de datos (ETL), ya que mediante este se realiza el análisis de la información generada en la empresa, la importancia del proceso ETL radica en el hecho de clasificar información para poder obtener datos relevantes contenidos en cada base de conocimiento, las dimensiones de este proceso permitirán a la gerencia la mejora de la toma de decisiones estratégicas que beneficien a la empresa, al personal, y al sector económico en el estado. Una de las ventajas que permitirá el proceso ETL es tener una visión global de los alcances obtenidos con respecto a proyectos exitosos anteriormente desarrollados es decir en base a los datos históricos. Otro aspecto que refleja una ventaja es facilitar la generación de estrategias de marketing para generar mejores oportunidades de negocio, mejorar las relaciones con los clientes, así como aumentar el número de ellos, es debido a esto que el proceso permitirá el éxito en la implementación de la tecnología de Business Intelligence.

### 1.3. Objetivo general

Por medio del proceso ETL se pretende consolidar la información para extraer conocimiento y datos, que estos a su vez sean aptos para alimentar al almacén de datos (Data Warehouse) y para ello también se pretende diseñar indicadores clave de rendimiento (KPIS) que aporten la orientación exacta para guiar al proceso y que esté pueda funcionar de manera adecuada, también se pretende realizar una simulación de lo que sería el proceso de ETL ya que para la implementación precisa se requiere de una alta infraestructura que puede requerir de más información de la que se tiene, en el capítulo siguiente sección 2.2 herramientas ETL y 2.3 tabla comparativa, se presenta un análisis a profundidad de algunas herramientas para el proceso.

### 1.4. Objetivos específicos

- Proponer KPIS para el control de proyectos
- Proponer mejoras en los reportes bajo la norma CMMI Dev 2 en caso de ser necesarias.
- Hacer propuestas para complementar información faltante en CMMI Dev 2.
- Obtener datos relevantes de los documentos bajo CMMI Dev 2.
- Obtener información para dar solución a los KPIS (indicadores)
- Hacer la adaptación del proceso ETL idóneo para la compañía.
- Explorar algunas opciones para realizar el proceso ETL.

### 1.5. Pregunta de investigación

¿Es suficiente la documentación que se maneja en CMMI Dev 2 para obtener información relevante que requieren las KPIS para así contribuir al proceso ETL y generar el conocimiento idóneo que aporte a la inteligencia de negocios?

## **1.6. Descripción del documento**

La presente tesis está conformada por seis capítulos. El primer capítulo Introducción adentra al lector para presentar el panorama general de la problemática principal del proyecto, justificación y objetivos establecidos. Para continuar en el capítulo 2 se encuentra la sección de términos relacionados con el tema principal ETL en Inteligencia de negocios, posteriormente se adentra al lector en el análisis de herramientas que permiten el procesamiento de información. En el capítulo 3 se presenta información esencial para el conocimiento de la norma de calidad implantadas en la empresa; los pasos esenciales y la dificultad del proceso son explicados en el capítulo 4, las pruebas del procesamiento de información son mostradas y explicadas en el capítulo 5, para finalmente presentar las conclusiones y los trabajos futuros en el capítulo 6.

## **1.7. Conclusión capítulo 1**

Gracias a la observación que se llevó a cabo en la organización se obtuvo el panorama general del estado inicial en la organización, lo que determinó la problemática principal y con ello se determinó el nivel de importancia del proyecto, los objetivos principales y el estado en el que se encuentra la empresa para solventar el proyecto, siempre con el fin de contribuir en la mejora de operaciones finales para el personal a cargo de la principal área que es fábrica de software y por lo tanto el reto más relevante para el tratamiento de la información.

## 2 Marco teórico

### 2.1. Conceptos

- Business intelligence (BI): se refiere a tecnologías, aplicaciones y prácticas para la recopilación, integración, análisis y presentación de la información comercial. El propósito de Inteligencia de Negocios es apoyar una mejor toma de decisiones empresariales. En esencia, los sistemas de Business Intelligence son Decision Support Systems basadas en datos (DSS). Inteligencia de Negocios a veces se usa de manera intercambiable con los libros de información, informe y herramientas de consulta y sistemas de información ejecutiva. (Olap.com, 2016)
- CMMi: CMMI es un modelo de mejora de desempeño de clase mundial para las organizaciones competitivas que quieren lograr operaciones de alto rendimiento. Demostrado ser eficaz en las organizaciones y los gobiernos a nivel mundial durante los últimos 25 años, CMMI consiste en las mejores prácticas recogidas diseñados para promover los comportamientos que conducen a un mejor desempeño en cualquier organización. Modelos CMMI Instituto contribuyen a identificar y mejorar las capacidades clave que elevan de su organización el rendimiento, la calidad y la rentabilidad. CMMI ofrece cuatro modelos que se pueden personalizar para satisfacer sus necesidades para diferentes entornos. (CMMI Institute, 2017)
- CMMI DEV 2: CMMI para desarrollo es un modelo de referencia que cubre las actividades del desarrollo y del mantenimiento aplicadas tanto a los productos como a los servicios. Las organizaciones de numerosas industrias, incluyendo la

aeroespacial, los bancos, la construcción de ordenadores, el software, la defensa, la fabricación del automóvil y las telecomunicaciones, utilizan CMMI para desarrollo. Los modelos de la constelación del CMMI para desarrollo contienen prácticas que cubren la gestión de proyectos, la gestión de procesos, la ingeniería de sistemas, la ingeniería del hardware, la ingeniería de software y otros procesos de soporte utilizados en el desarrollo y el mantenimiento. El modelo CMMI para desarrollo + IPPD cubre también la utilización de equipos integrados que están implicados en las actividades de desarrollo y mantenimiento (IPPD).

- Dataware house: es un repositorio unificado para todos los datos que recogen los diversos sistemas de una empresa. El repositorio puede ser físico o lógico y hace hincapié en la captura de datos de diversas fuentes sobre todo para fines analíticos y de acceso. Normalmente, un data warehouse se aloja en un servidor corporativo o, en la nube. Data Warehouse es una arquitectura de almacenamiento de datos que permite a los ejecutivos de negocios organizar, comprender y utilizar sus datos para tomar decisiones estratégicas. (powerdata, 2015)
- ETL: este término viene de inglés de las siglas Extract-Transform-Load que significan Extraer, Transformar y Cargar y se refiere a los datos en una empresa. ETL es el proceso que organiza el flujo de los datos entre diferentes sistemas en una organización y aporta los métodos y herramientas necesarias para mover datos desde múltiples fuentes a un almacén de datos, reformatearlos, limpiarlos y cargarlos en otra base de datos, data mart ó dataware house. Forma parte de la Inteligencia Empresarial (Business Intelligence), también llamado “Gestión de los Datos” (Data Management). ETL lee los datos primarios de unas bases de datos de sistemas principales, realiza la transformación, validación, el proceso cualitativo, filtración y al final escribe datos en el almacén y en este momento los datos son disponibles para analizar por los usuarios. (Guru, 2017)
- Key Performance Indicators (KPI): Un Indicador de Desempeño Clave es un

valor medible que demuestra cuán efectivamente una compañía está logrando los objetivos clave del negocio. Las organizaciones usan indicadores de desempeño en múltiples niveles para evaluar su éxito al alcanzar los objetivos. Los KPIs de alto nivel pueden enfocarse en el desempeño general de la empresa, mientras que los KPIs de bajo nivel pueden centrarse en procesos en departamentos como ventas, marketing o un centro de llamadas. (klipfolio, 2017)

- MoProSoft: es un modelo de procesos para la industria de software en México. Desarrollado por iniciativa de la Secretaría de Economía y con el apoyo de empresarios y académicos mexicanos. Fomenta la estandarización a través de buenas practica en la gestión y desarrollo de software. Establece y emplea un patrón para definir cada proceso. El patrón de procesos es una agrupación esquemática de los elementos que configuran un proceso. La implementación de MoProSoft mejorar la calidad del software desarrollado por la empresa. Pretende elevar la capacidad de las empresas para alcanzar niveles altos de calidad y aumentar así su competitividad. Este modelo permite a las empresas mexicanas medir su nivel de madurez.
- OLTP: OnLine Transaction Processing, y estos procedimientos facilitan las aplicaciones transaccionales para la entrada de datos, la recuperación y el procesamiento. Es típico de las bases de datos operacionales, y los paquetes de software utilizados para estas tecnologías están basados en la técnica cliente-servidor. Dichos sistemas los suelen utilizar empresas que disfrutan de una red informática distribuida. Este sistema optimiza el acceso a los datos, facilitando las tareas de lectura, escritura y análisis que se llevan a cabo de manera frecuente. Los procedimientos OLTP estructuran estos datos según su nivel de aplicación, además, no tienen por qué mantenerse uniformes en los diferentes departamentos por la común falta de compatibilidad. (Chain, 2016)

## 2.2. Herramientas ETL

### 2.2.1. Talend open studio para la integración de datos

Herramienta competitiva que ofrece gran cantidad de software para el procesamiento de información, contiene diferentes productos para las distintas necesidades del cliente. Talend ofrece la integración de datos robustos en una arquitectura abierta y escalable para maximizar el valor a su negocio. Aumenta la productividad con una robusta herramienta de código abierto ETL se puede descargar de forma gratuita y empezar a utilizar hoy en día. Integra fácilmente todos sus datos con su almacén de datos y aplicaciones, o sincroniza datos entre sistemas (talend, 2017)

Características:

- Software para el intercambio de metadatos.
- Herramienta para desarrolladores basadas en Eclipse Soporte de ETL.
- Exportación y ejecución de trabajos independientes en entornos de ejecución.
- SaaS: Marketo, Salesforce, NetSuite y más.
- Aplicaciones empaquetadas: SAP, Microsoft Dynamics, Sugar CRM y más.
- Tecnologías: Dropbox, Box, SMTP, FTP / SFTP, LDAP.
- Controlar y orquestar flujos de datos e integraciones de datos con trabajos maestros.
- Mapa, agregación, clasificación, enriquecimiento y fusión de datos.

### 2.2.2. ETL Informatica powercenter

Respalda todo el ciclo de vida de integración de datos, desde el impulso del primer proyecto hasta la garantía de implantaciones empresariales de misión crítica satisfactorias. PowerCenter constituye la base de todas sus iniciativas de integración de datos, como análisis y data warehousing, migración de aplicaciones o consolidación

y gobernanza de datos. (informatica, 2016)

Características:

- Colaboración entre negocio y TI.
- Creación de prototipos, validación y perfilados rápidos.
- Reutilización, automatización y facilidad de uso.
- Conectividad universal.
- Gestión basada en metadatos.
- Escalabilidad, rendimiento y nula inactividad del sistema.
- Pruebas de validación de datos automatizadas.
- Supervisión operacional y de gobierno.
- Transformación avanzada de datos.
- Datos en tiempo real para las aplicaciones y los análisis.

### 2.2.3. IBM Datastage

IBM InfoSphere DataStage es la plataforma ETL de primer nivel para integrar datos confiables a través de múltiples sistemas empresariales que aprovechan un marco paralelo de alto rendimiento en las instalaciones o en la nube. DataStage apoya la gestión de metadatos extendido y conectividad empresarial omnipresente. Esta potente plataforma escalable proporciona integración flexible de datos heterogéneos, incluyendo grandes datos en reposo (Hadoop-base) o grandes datos en movimiento (basado-stream), en ambas plataformas distribuidos y de mainframe. (IBM, 2015-2017)

Características:

- Soporta grandes volúmenes de datos y Hadoop.

- Mejora la eficiencia de la empresa ETL.
- Ofrece una potente plataforma ETL.
- Proporcionar datos ETL confianza en cualquier momento y en cualquier lugar.
- Resolver problemas complejos de datos grandes.
- Añade funciones sin costo adicional.
- Integrar fácilmente aplicaciones de la nube.

### 2.2.4. Oracle data Integrator

Proporciona una productividad de desarrollador superior y una experiencia de usuario mejorada con una rediseñada interfaz de usuario declarativa basada en flujo y una integración más profunda con Oracle GoldenGate. ODI12c se basa más en su arquitectura flexible y de alto rendimiento con un gran soporte de datos y agrega paralelismo a la hora de ejecutar procesos de integración de datos. (Oracle , 2009)

Características:

- Capacidad de manejar grandes volúmenes de datos con un desempeño óptimo para cargar Data Warehouse y Data Mart.
- Maneja cargas incrementales, integridad de datos, reglas de negocio y consistencia.
- Funcionalidad de invocar servicios externos para propósitos de integración e implementar servicios de integración y transformación integrados a una arquitectura orientada a servicios.
- Master Data Management, combina aplicaciones y tecnologías que consolidan, limpian, mejora los datos maestros de la empresa y los sincroniza con aplicaciones, procesos de negocio y herramientas analíticas como Oracle.
- Migración, provee cargas masivas eficientemente de datos históricos, incluyendo transformaciones complejas de sistemas heredados a sistemas nuevos.

### 2.3. Tabla comparativa de herramientas para el proceso ETL

La tabla 2.1 representa una evaluación de las herramientas presentadas anteriormente, considerando las ventajas y desventajas que proporcionan específicamente para los procesos involucrados en ETL.

**Tabla 2.1:** Comparación de herramientas

Herramientas	Ventajas	Desventajas
Talend Open Studio	1.Herramienta Open Source. 2. Es compatible con cualquier. plataforma java o perl. 3. Utiliza poca capacidad de Ram. 4. Cuenta con herramientas de monitoreo y registro histórico.	1. Es lento comparado con pentaho. 2. Carece de soporte.
ETL Informática PowerCenter	1. Es veloz gracias a PushDown. 2. Fácil de utilizar. 3. Ofrece soporte mundial.	1. Existe un nivel medio de riesgo. 2. Requiere de capacitación. 3. Para mejorar la calidad de datos requiere del complemento Data Quality.
IBM DataStage	1. Es muy rápido para DB2. 2. Para la facilidad de uso requiere de un entorno GUI. 3. Ofrece soporte mundial si se contrata un paquete.	1. Disminuye la velocidad con otros gestores de BD. 2. Carece de monitoreo. 3. Requiere de otro complemento IBM DATA manager. 4. Para mejora la calidad de datos requiere Cogno Data Manager.
Oracle Data Integrator	1. La velocidad es proporcional al servicio de Oracle en el que se esté trabajando. 2. Tiene facilidad de uso para BD Oracle debido a Data Pump. 3. Ofrece soporte local para Latinoamérica.	1. Requiere de complementos Oracle GoldenGate para la Integración y Replicación de Datos en Tiempo Real. 2. No existe mucha compatibilidad con otros tipos de bases de datos. 3. Mejora la calidad de datos a través de Oracle Warehouse Builder Data profiling features.

## 2.4. R Software libre

R es una herramienta de software libre para el análisis estadístico, es considerado uno de los mejores ya que aporta una gran variedad de métodos estadísticos para generar gráficos que aporten al análisis de datos. Su funcionamiento se basa en la programación a base de comandos lo cual representa una ventaja y desventaja para algunos de los usuarios, pero ante esta situación existe una paquetería llamada R commander que a decir por la comunidad de R permite utilizarlo sin tener que escribir comandos y solo con la utilización del ratón.

El software permite el acceso a materiales de aprendizaje, manipular ficheros de datos, realizar representaciones gráficas y realizar análisis estadístico de gran complejidad o de menor.

R fue desarrollado por dos jóvenes profesores de estadística Robert Gentleman y Ross Ihaka en 1995 con su primera versión sin embargo ha crecido debido al aporte de sus diferentes colaboradores permitiendo así el manejo a distintos investigadores, es por ello que representa uno de los lenguajes más utilizados, no solo del sector de estadística sino también en la minería de datos, la investigación biomédica, bioinformática, matemáticas financieras entre otros es por esto que tiene diversidad de paqueterías con diferentes funcionalidades.

Dentro de las opciones para la exploración de mejores herramientas para el proceso ETL está R, debido a que ofrece diferentes paqueterías que pueden aportar una solución para el adecuado procesamiento de información, sin embargo, representa una desventaja el tener que desarrollar una herramienta mediante esta plataforma ya que requiere de la inversión de tiempo para programar algunas funcionalidades que ya pueden estar diseñadas en alguna herramienta específica de ETL como las presentadas con anterioridad.

## 2.5. Estado del arte: proceso ETL y key performance indicators

El proceso de extracción, transformación y carga de datos es una de las tareas fundamentales para la implementación de inteligencia de negocios en cualquier sector, es por ello por lo que no es algo completamente nuevo ya que si se ha tenido éxito en la implementación de esta tecnología es debido al aporte que se requiere, existen trabajos que buscan la mejor propuesta o adaptación para la implementación de este proceso considerando que se debe llevar a cabo en base a las necesidades que requiere cada tipo de proyecto u organización. A continuación, se plantean algunos trabajos relacionados que involucran distintos aspectos interesantes para el aporte al proyecto.

### Artículos de ETL

(Antolić, 2006) define KPI's para la evaluación de la eficiencia de procesos implicados en el desarrollo de software; mediante evaluaciones comparativas se obtienen registros de los proyectos desarrollados para posteriormente ser comparados y así buscar las mejores prácticas en el desarrollo de proyectos. La observación en general que se plantea es la recopilación y análisis de estos datos que fueron hechas mensualmente y se hace énfasis en la evolución que deben tener los KPIS en base a las necesidades de información para que esto conlleve a acciones de mejora continua.

(Tang Jun; Cui Kai; Feng Yu; Tong Gang, 2009) articulo en el que los autores proponen la reutilización de software para codificación ETL, y para ello utilizan la tecnología JavaBean, la tecnología JDBC y la tecnología de procedimiento almacenado de llamada JSP, para elaborar sobre la función de la reutilización de software en el proceso ETL [13].

(Tang Jun; Cui Kai; Feng Yu; Tong Gang, 2009) Comparan dos tipos de desarrollo ETL: el primero mediante el uso de herramientas ETL y el segundo mediante la codificación manual; Mediante la primera se obtiene un desarrollo con menos tiempo y reduce la complejidad, la codificación manual ofrece la ventaja de la flexibilidad

y la eficiencia sin embargo en ella existe un alta demandad de asistencia técnica. En este artículo se utiliza el modelo de arquitectura de tres niveles J2EE. Clinte: aplicación bajo el lenguaje JAVA, JDBC como servidor de aplicaciones y sistema de administración de base de datos ORACLE.

(Darshan M., Amit , Y P , C.K, 2010) reflejan la importancia de la creación de un almacén de datos que sea capaz de detectar los cambios dentro de los datos ya almacenados para evitar que el sistema se vuelva lento es por ello que proponen dos pasos: el primero es unir y el segundo agregación, estos evitaran la demora en la actualización del almacenamiento mediante lo que denominan CDC (cambio de captura de datos) mediante el procesamiento de los cambios se realiza la integración de los datos por lo tanto extraer se vuelve una tarea más eficiente y reduce la latencia es decir replica los cambios en los sistemas fuente y se por lo tanto se pone a disposición del usuario en el depósito de datos.

(Vilanueva, Chavez, Li, 2011) proponen un proceso ETL basado en dos componentes principales, una metodología para describir en detalle los pasos del proceso y una biblioteca de componentes de software para ayudar al proceso. La función principal de la ontología es modelar los elementos de dominio del almacén de datos, conceptos y relaciones. El proceso ETL en este articulo requiere de los siguientes pasos: extracción de metadatos, adición de metainformación y generación del modelo de datos lógicos; es importante resaltar que para la implementación de esta solución se consideraron unicamente dos fuentes de información: archivos XML y bases de datos.

(Bustamante Martínez, Amaru Galvis, Gómez Flórez, 2012) articulo dedicado a la investigación de los cuatro tipos de modelado ETL: las inspiradas en el paradigma de flujo y procesos, las inspiradas en el paradigma de programación orientada a objetos, y los paradigmas UML. También comparten la experiencia de desarrollar con la técnica de modelado UML, sin embargo destacan que ninguna de ellas es considerada como

una buena práctica o un estándar y por ello recomiendan el uso de UML combinado con otra tecnología.

(Bansal, 2014) documenta la experiencia del manejo del proceso ETL para bigData y el gran reto que es detectar información de relevancia dentro de grandes volúmenes de datos, propone la creación de un modelo de datos semánticos para integrar y comprender datos de múltiples fuentes que requieran de extracción para generar conocimiento.

(Hamed, 2015) resalta la importancia del proceso para detectar datos de calidad y la complejidad del análisis, diseño y desarrollo de aplicaciones orientadas a estos datos. A su vez se plantea que la complejidad del proceso ETL se debe a las actividades de: integrar eficazmente sistemas que tengan diferentes sistemas de gestión de base de datos, sistemas operativos, hardware y protocolos de comunicación. La autora propone un modelado de conocimiento y hace referencia a algunos criterios para la consideración de calidad de datos como: complejidad, coherencia, validez, conformidad, precisión y puntualidad.

(Hamizah, Mohd, Lizawati, 2015) los autores utilizan el proceso ETL para la generación de inteligencia de negocios para una universidad, de esta forma detectar las oportunidades y los riesgos en base a sus múltiples fuentes de información, hacen hincapié en la integración de datos y las herramientas de ETL.

(Spalevic, Emir, Ilic, Mladen, Stefan, 2016) hacen uso de la herramienta Talend Open Studio como medio para la realización del proceso ETL y se expone como propuesta principal la expansión de la estructura básica del proceso ETL, a su vez mencionan que de esta manera se pueden resolver problemas en la fase de carga, transformación o almacenamiento. Los autores argumentan que con esta idea se puede obtener un mejor control sobre los sistemas remotos y al implementarlo correctamente

se obtiene un registro completo de todo el proceso de ETL asociado al análisis de datos que hace el DW.

(Chen, Baoran An, Yan, 2016) proponen un nuevo marco de flujo de trabajo para la ejecución paralela de ETL basado en un sistema multi-agente. En este trabajo se plantea el problema que existe para la generación de sistemas de inteligencia de negocios enfocándose totalmente al proceso ETL, debido a que durante este proceso se detecta la generación de datos de múltiples tipos, formas complejas y de gran cantidad. Por lo anterior es importante considerar que la generación de datos se da en tiempo real, alta velocidad y de forma continua. A causa de ello los autores proponen un sistema extendido de ETL tradicional basado en agentes múltiples para soportar la ejecución paralela de ETL.

(Jianwen Yan, 2017) presentan un nuevo marco para análisis de negocios basado en registros de flujo de trabajo, los autores formulan un nuevo marco para análisis de negocios basado en registros de flujo de trabajo, hacen uso de registros de registros para detectar los datos a los que se accede y se modifican. Los flujos de trabajo analizan claramente los problemas analíticos a partir de detalles irrelevantes del sistema y permiten que las técnicas de BI sean aplicadas generalmente, por lo tanto existe una mejora con respecto al tiempo en ejecución del proceso.

(Loreto, y otros, 2017) hacen uso de la inteligencia de negocios e indicadores para ofrecer ventajas en la atención de salud así como para medir la calidad del servicio, también incluyen su experiencia con la suite de herramientas para BI, dentro de la investigación que llevaron a cabo detectaron inconsistencias que son generadas por la extracción manual y que dan como consecuencia al retraso de la toma de decisiones en tiempo real, sin embargo, proponen la modificación del proceso ETL para que la extracción se realice de forma automática.

(Adnan, Ilham, Usman, 2017) es un artículo que tiene como objetivo aplicar Hadoop para procesar datos en cada etapa ETL, presentar la solución de código abierto de extremo a extremo y con una herramienta ETL, utilizando Ethernet adaptan iSCSI como almacenamiento y conectandolo a la red (NAS) y para que pueda fungir como medio de almacenamiento alternativo, esta propuesta es para las pyme de TI que desean crear un almacén de datos eficiente sin la necesidad de invertir grandes recursos. Al final de la implementación los autores concluyen que existe un impacto relevante en la localidad en la que se desarrolla el proceso, por lo tanto, existe una variación en los resultados cuando se utiliza un almacenamiento remoto y cuando se realiza de manera local, el tiempo para la extracción es favorable en el almacenamiento NAS más que en el local y la transformación muestra resultados cambiantes en ambos.

(Chu, Kuo, Cheng, 2017) proponen una aplicación web basada en la paquetería de R studio Shiny con el fin de crear aplicaciones que sirvan como herramientas ETL de metadatos que ayuden a la limpieza, análisis e informe de datos de salud y de esta manera revelar la mejoría en las terapias para el tratamiento del cáncer, sin embargo no existe una aplicación para esta tecnología.

## 2.6. Conclusión capítulo 2

Gracias a este capítulo se permite conocer algunas de las terminologías que involucran a la inteligencia de negocios ya que el estudio de las distintas herramientas permite determinar la mejor alternativa para llevar a cabo una implementación eficiente del ETL. Mediante los trabajos relacionados que fueron abstraídos de una de las principales fuentes de información para los investigadores, se permite determinar la metodología más conveniente o de lo contrario realizar ajustes que puedan adaptarse a la organización y con ello proponer una solución acorde a sus necesidades, por lo tanto, fue de suma importancia adquirir conocimiento de los proyectos ETL anteriormente realizados.

## 3 Metodología

En este capítulo se presenta información que permitirá comprender la importancia que tiene la norma con la que se encuentra certificada la empresa, con un enfoque principal al nivel 2 del modelo. Así mismo se presentan herramientas que dan la oportunidad de conocer el impacto que generan cada uno de los objetivos que define CMMI y el impacto que tiene para lograr la aplicación de inteligencia de negocios dentro de la cual tiene como paso principal el proceso ETL y de qué manera aporta información para dar una solución eficiente a las KPI'S.

### 3.1. Modelo de madurez de capacidad integrado (CMMI)

Fue creado por el instituto de ingeniería de Software para ayudar a las empresas a dar un enfoque en base a la mejora de los procesos y garantizar que los procesos involucrados en el desarrollo de productos o servicios se lleve a cabo adecuadamente. (Mary Beth Chrissis, 2009) CMMI es útil para el personal, cliente y directivos debido a que permite conocer el nivel de madurez en el que se encuentra la empresa en base a la evaluación de los procesos que son los encargados de producir software.

CMMI para el desarrollo de software considera las buenas prácticas conforme a las actividades de mantenimiento y creación que deben ser ajustadas a los productos o servicios que se ofrecen, para así lograr la entrega de buenos productos y mediante ello lograr la satisfacción tan deseada del cliente. Los procesos que se definen dentro

de la norma de calidad son aquellos que utilizan las organizaciones, debido a que permiten consolidar la forma de operar y a su vez evolucionar para incorporar los nuevos conocimientos de mejora que se adquieren con la experiencia y mediante ello saber que recursos se poseen.

CMMi trabaja para lograr procesos más eficientes y eficaces que permitan contribuir al logro de objetivos estratégicos, obteniendo un trabajo más inteligente y por tanto remunerable. A partir de la versión 1.2 de este modelo se presenta el concepto de constelaciones. Una constelación es el conjunto de componentes que se utiliza para construir modelos, así como materiales para capacitaciones y también para poder evaluar las áreas de interés. La figura 3.1 representa un diagrama en el cual se representan los 3 tipos de constelaciones definidos. (Mary Beth Chrissis, 2009)

<b>Desarrollo (DEV)</b>	<b>Adquisicion (ACQ)</b>	<b>Servicios (SVC)</b>
<ul style="list-style-type: none"> <li>•Esta guía permite medir, monitorear y controlar el proceso de desarrollo y mantenimiento de Software</li> </ul>	<ul style="list-style-type: none"> <li>•Es la guía establecida para la mejora del proceso de adquisicion de productos y servicios.</li> </ul>	<ul style="list-style-type: none"> <li>•Esta guía abarca las actividades que se requieren para gestionar, establecer y entregar servicios.</li> </ul>

**Figura 3.1:** Constelaciones de CMMI.

La constelación con la que se trabaja en esta tesis es la de desarrollo (CMMI Dev) debido a que la empresa trabaja con dicha constelación, diversas organizaciones trabajan con la versión de desarrollo de CMMI entre ellas están: aeroespacial, bancos, construcción de computadoras, construcción de software, la fabricación de automóviles y las telecomunicaciones. Esta constelación fue diseñada para abarcar la gestión de proyectos, gestión de procesos, ingeniería de sistemas, ingeniería de hardware, ingeniería de software y el proceso de soporte para mantenimiento.

### 3.1.1. Componentes del modelo

El modelo abarca diferentes estrategias para el control de actividades por lo tanto se agrupa en tres categorías, a continuación, se explica el fin con el que fueron establecidos:

- Componentes requeridos: establecen los lineamientos con los que debe cumplir la empresa, con el fin de cumplir los objetivos de un área de proceso específica.
- Componentes esperados: describen las actividades que pueden ser llevadas a cabo para obtener un componente requerido.
- Componentes informativos: contienen los detalles que sirven para considerar las acciones que se deben llevar a cabo para obtener los componentes requeridos y los esperados.

Las áreas de proceso son las acciones conjuntas que deben llevarse a cabo durante el desarrollo de un proyecto, este conjunto de actividades lleva al cumplimiento de los objetivos de mejora de proceso para cada área, el modelo establece 22 áreas de proceso se pueden observar en la figura 3.2 así como su nivel de madurez.

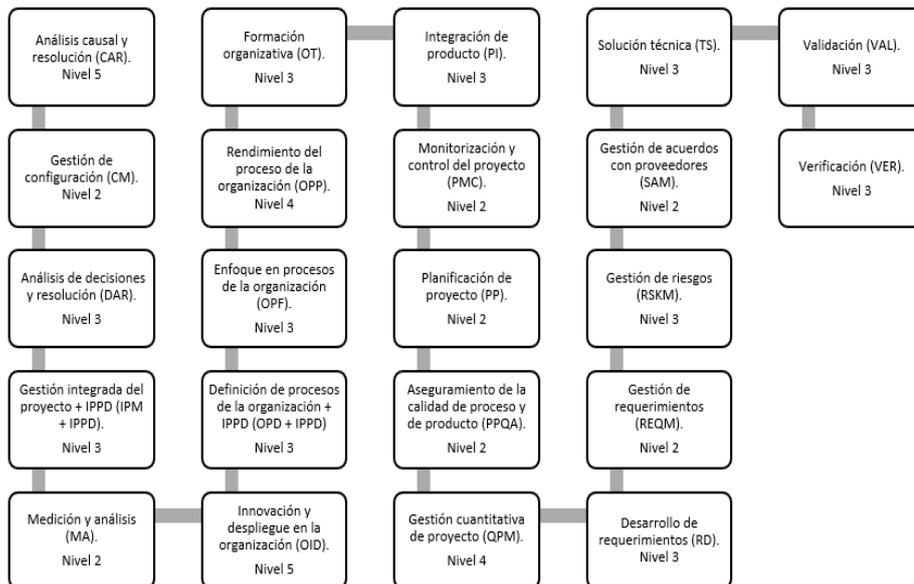


Figura 3.2: Áreas de Proceso CMMI DEV.

CMMI establece el nivel de madurez por etapas, permite evaluar a cada proceso asignando a cada uno un grado del 1 al 5 en base al cumplimiento de los objetivos. El nivel de madurez representa el valor de evolución que tienen los procesos establecidos en la empresa, cada nivel es conformado por diferentes áreas de proceso y se pueden medir mediante el cumplimiento de las metas genéricas y específicas, los puntos enumerados a continuación son los niveles de madurez por etapas:

1. Inicial.
2. Gestionado.
3. Definido.
4. Gestionado cuantitativamente.
5. En optimización.

### **3.1.2. Nivel de madurez 2 Gestionado**

Cuando una empresa cuenta con esta certificación se sabe que los procesos de la organización son planificados y llevados a cabo de acuerdo con lo establecido, a su vez son monitoreados, controlados y supervisados para detectar defectos que puedan presentarse durante el desarrollo cotidiano de las actividades. Se dice que una empresa cuenta con un buen nivel de calidad en sus procesos cuando se presentan defectos o imprevistos y aun así mantiene la calidad con la que operan las áreas de proceso, por lo tanto, mantiene la transparencia en la entrega de hitos (etapas del proyecto) ante las partes interesadas del proyecto. Este nivel es destacado debido a que distribuye las actividades entre los miembros del equipo para ganar el compromiso del personal al revisar las actividades que les fueron otorgadas, aunado a ello que el producto final logre la satisfacción de las necesidades del cliente, así como las que establece el proceso y los estándares de calidad. La figura 3.3 representa los procesos que abarca el nivel de madurez 2 de CMMI.



Figura 3.3: Áreas de proceso del nivel 2.

### 3.1.3. Categoría de las áreas de proceso

En CMMI las áreas de proceso están clasificadas mediante cuatro categorías: gestión de procesos, gestión de proyectos, ingeniería y soporte; estas contienen las áreas de proceso enfocadas al cumplimiento de la categoría, pero existe interacción continua entre ellas. La figura 3.4 muestra la clasificación de las cuatro categorías e incluye los nombres de los procesos únicamente del nivel 2 que son las dos categorías que abarca CMMI en este nivel de madurez.

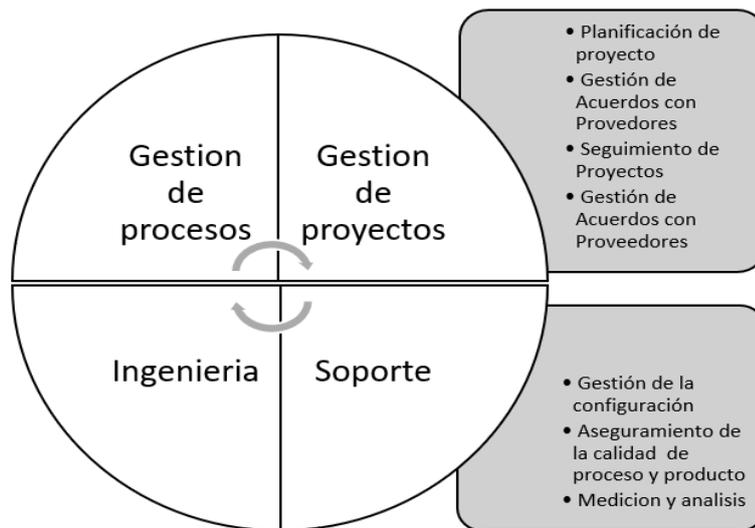


Figura 3.4: Clasificación de las áreas de proceso del nivel 2 en base a su categoría.

### 3.1.4. Objetivos de las Áreas de proceso Nivel 2

A continuación, se presentan las siete áreas de proceso de CMMI para desarrollo del nivel 2 adjuntando el propósito, las metas y practicas específicas para las que fueron diseñadas.

### *Planificación de proyecto*

El propósito de la Planificación de proyecto (PP) es establecer y mantener planes que definan las actividades del proyecto.

Metas y prácticas específicas.

- SG 1 Establecer estimaciones.
  - SP 1.1 Estimar el alcance del proyecto.
  - SP 1.2 Establecer las estimaciones de los atributos del producto de trabajo y de las tareas.
  - SP 1.3 Definir el ciclo de vida del proyecto.
  - SP 1.4 Determinar las estimaciones de esfuerzo y de coste.
  
- SG 2 Desarrollar un plan de proyecto.
  - SP 2.1 Establecer el presupuesto y el calendario.
  - SP 2.2 Identificar los riesgos del proyecto.
  - SP 2.3 Planificar la gestión de los datos.
  - SP 2.4 Planificar los recursos del proyecto.
  - SP 2.5 Planificar el conocimiento y las habilidades necesarias.
  - SP 2.6 Planificar la involucración de las partes interesadas.
  - SP 2.7 Establecer el plan de proyecto.
  
- SG 3 Obtener el compromiso con el plan.
  - SP 3.1 Revisar los planes que afectan al proyecto.
  - SP 3.2 Reconciliar los niveles de trabajo y de recursos.
  - SP 3.3 Obtener el compromiso con el plan.

### *Seguimiento de proyectos*

El propósito del seguimiento de proyectos PMC es Monitorear las actividades del

proyecto y tomar acciones correctivas documentadas.

Metas y prácticas específicas.

- SG1 Monitorear el desarrollo del proyecto contra el plan de proyecto.
  - SP 1.1 Monitorizar los parámetros de planificación del proyecto.
  - SP 1.2 Monitorizar los compromisos.
  - SP 1.3 Monitorizar los riesgos del proyecto.
  - SP 1.4 Monitorizar la gestión de los datos.
  - SP 1.5 Monitorizar el involucramiento de las partes interesadas.
  - SP 1.6 Llevar a cabo las revisiones del progreso.
  - SP 1.7 Llevar a cabo las revisiones de hitos.
  
- SG2 Administrar acciones correctivas hasta el cierre.
  - SP 2.1 Analizar las cuestiones.
  - SP 2.2. Llevar a cabo las acciones correctivas.
  - SP 2.3 Gestionar las acciones correctivas hasta su cierre.

#### *Administración de requerimientos*

El propósito de la administración de requerimientos (REQM) es gestionar los requerimientos de los productos y de los componentes del producto del proyecto, e identificar inconsistencias entre esos requerimientos, los planes y productos de trabajo del proyecto.

Metas y prácticas específicas

- SG 1 Gestionar los requerimientos.
  - SP 1.1 Obtener una comprensión de los requerimientos.
  - SP 1.2 Obtener el compromiso sobre los requerimientos.
  - SP 1.3 Gestionar los cambios de los requerimientos.

- SP 1.4 Mantener la trazabilidad bidireccional de los requerimientos.
- SP 1.5 Identificar las inconsistencias entre el trabajo del proyecto y los requerimientos.

### *Gestión de acuerdos con proveedores*

El propósito de la Gestión de acuerdos con proveedores (SAM) es gestionar la compra de productos.

Metas y prácticas específicas

- SG 1 Establecer los acuerdos con proveedores.
  - SP 1.1 Determinar el tipo de compra.
  - SP 1.2 Seleccionar los proveedores.
  - SP 1.3 Establecer los acuerdos con el proveedor.
- SG 2. Satisfacer los acuerdos del proveedor.
  - SP 2.1 Realizar el acuerdo del proveedor.
  - SP 2.2 Monitorizar los procesos seleccionados del proveedor.
  - SP 2.3 Evaluar los productos de trabajo seleccionados del proveedor.
  - SP 2.4 Aceptar los productos adquiridos.
  - SP 2.5 Transferir los productos.

### *Gestión de configuración*

El propósito de la Gestión de configuración (CM) es establecer y mantener la integridad de los productos de trabajo utilizando la identificación de configuración, el control de configuración, el registro del estado de configuración y las auditorías de configuración.

Metas específicas y prácticas específicas

- SG 1 Establecer líneas base.
  - SP 1.1 Identificar elementos de configuración.
  - SP 1.2 Establecer un sistema de gestión de configuración.

- SP 1.3 Crear o liberar líneas base.
- SG 2 Seguir y controlar los cambios.
  - SP 2.1 Seguir las peticiones de cambio.
  - SP 2.2 Controlar los elementos de configuración.
- SG 3 Establecer la integridad.
  - SP 3.1 Establecer registros de gestión de configuración.
  - SP 3.2 Realizar auditorías de configuración.

*Aseguramiento de la calidad de proceso y de producto*

El propósito de Aseguramiento de la calidad de proceso y de producto (PPQA) es proporcionar al personal y a la gerencia una visión objetiva de los procesos y de los productos de trabajo asociados.

Metas y prácticas específicas

- SG 1 Evaluar objetivamente los procesos y los productos de trabajo.
  - SP 1.1 Evaluar objetivamente los procesos.
  - SP 1.2 Evaluar objetivamente los productos de trabajo y los servicios.
- SG 2 Proporcionar una visión objetiva.
  - SP 2.1 Comunicar y asegurar la resolución de las no conformidades.
  - SP 2.2 Establecer registros.

*Medición y análisis*

El propósito de Medición y Análisis (MA) es desarrollar y mantener la capacidad de medición utilizada para dar soporte a las necesidades de información de la gerencia.

Metas y prácticas específicas

- SG 1 Alinear las actividades de medición y análisis.
  - SP 1.1 Establecer los objetivos de medición.

- SP 1.2 Especificar las medidas.
- SP 1.3 Especificar los procedimientos de recogida y de almacenamiento de datos.
- SP 1.4 Especificar los procedimientos de análisis.
- SG 2 Proporcionar los resultados de la medición.
  - SP 2.1 Recoger los datos de la medición.
  - SP 2.2 Analizar los datos de la medición.
  - SP 2.3 Almacenar los datos y los resultados.
  - SP 2.4 Comunicar los resultados.

En base a los criterios establecidos en la sección anterior las figuras 3.5, 3.6 y 3.7 muestran algunas capturas de pantalla de los documentos generados en la organización, con los que se da el cumplimiento a los objetivos de la norma de calidad.

*Área de proceso: Planificación de proyecto.*

*Meta: Establecer estimaciones.*

*Práctica específica: Estimar el alcance del proyecto.*

**4. Alcance del proyecto**

Desarrollar el Sistema de Solicitudes de Viáticos de acuerdo a las especificaciones

El Sistema de Solicitudes de Viáticos, incluye los siguientes módulos:

1. Inicio de sesión
2. Viáticos (web y móvil)
  - a) Solicitud de Viático
  - b) Comprobación de Gastos
3. Verificación
  - a) Verificación de Solicitudes
  - b) Verificación de Gastos
4. Aprobación de Solicitud de Viáticos
  - a) Concentrado de Solicitudes
  - b) Realización de Pagos
5. Validación de Viáticos (web y móvil)
6. Configuraciones
  - a) Tabuladores de viáticos
  - b) Tabuladores de comprobantes

4.1 Distribución de esfuerzo (WBS para control de horas)

Fecha de actualización		08/09/2014		
Etapas/Fase	Actividad	Subactividad	Esfuerzo planeado	Esfuerzo horas
Planeación Inicial	Elaborar Plan General		20.0	20.0
	Entrevista inicial con el cliente		4.0	4.0
	Definición y solicitud de repositorio		4.0	4.0
	Realizar Plan de Desarrollo		4.0	4.0
	Reunion equipo de trabajo		4.0	4.0
Especificación de Requerimientos			152.0	163.0
Realización de Requerimientos	Realizar Entrevistas (En sitio con el cliente)		27.0	29.0
	Elaborar BPMNs		6.0	6.0
	Elaborar Lista de Requerimientos y especificación de requerimientos		43.0	46.0
	Elaborar Matriz de Trazabilidad		8.0	9.0
	Elaborar Diseño Preliminar de Pantallas		15.0	18.0
	Elaborar Plan de Pruebas		18.0	19.0
Plan Integral del Proyecto	Elaborar Manual de Usuario Preliminar		14.0	14.0
	Elaborar tareas de seguimiento de fase		21.0	22.0
Plan Integral del Proyecto	Elaboración de descripción de proyecto y registro de proyecto		41	44
	Tiempo Estimado		5.0	5.0
			6.0	7.0

**Figura 3.5:** Documentación para estimar el alcance del proyecto.

Área de proceso: Gestión de configuración.

Meta: Seguir y controlar los cambios.

Práctica específica: Seguir las peticiones de cambio.

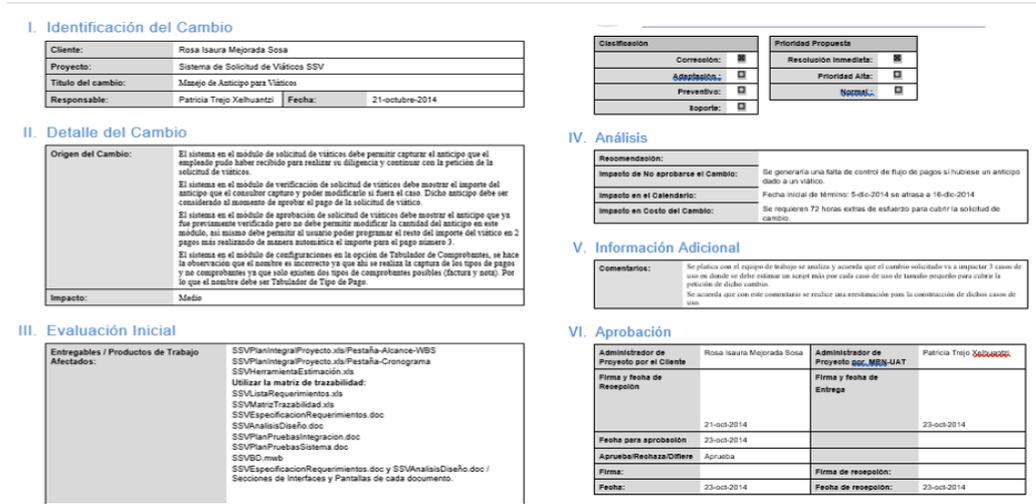


Figura 3.6: Documentación para el seguimiento de cambios.

Área de proceso: Aseguramiento de la calidad de proceso y de producto

Meta: Evaluar objetivamente los procesos y los productos de trabajo

Práctica específica: Evaluar objetivamente los productos de trabajo y los servicios



Figura 3.7: Documentación para la evolución del producto.

### 3.2. Key performance Indicators

Key performance Indicators o en español indicador clave de desempeño es un tipo de medida que evalúa el rendimiento que tienen los procesos, estos deben definirse en base a lo que se desea medir y deben ser enfocados por completo a un proceso en específico. Esta medida de desempeño puede ser cuantificable o cualitativa, es decir puede brindarnos información tanto para finanzas como también para el control de proyectos. Mediante ellas se puede conocer la rentabilidad que tiene una organización y determinar el tipo de acciones que deben llevarse a cabo para mejorar los aspectos con mayor carencia de eficiencia.

(Parmenter, 2007) Define siete características para los indicadores:

1. Medidas no financieras que no involucran ningún tipo de moneda. En algunas ocasiones se desarrollan de forma errónea, es decir: “cuando colocan un signo de dólar en una medida y se cree que ya se ha convertido en un indicador de resultado”, por ejemplo: las ventas diarias son el resultado de las actividades que se han llevado a cabo para crear las ventas. Sin embargo, los indicadores no necesariamente son medidas que involucren actividades monetarias.
2. Medido con frecuencia (por ejemplo, diariamente o 24/7) Una medida mensual, trimestral o anual no puede ser un KPI, debido a que no son clave para el negocio, los KPI son medidas actuales o futuras en lugar de medidas pasadas, por ejemplo: el número de visitas de clientes clave planificadas el próximo mes o una lista por cliente clave de la fecha de la próxima visita planificada. Este ejemplo ayuda a comprender el punto principal ya que son indicadores pasados que miden los eventos del último mes o trimestre. Por lo tanto, estos indicadores no pueden ser indicadores clave de desempeño.
3. Actuado por el CEO y el equipo directivo Se dice que los buenos KPI hacen la diferencia ya que involucran de manera continua al personal conformado por el CEO o director ejecutivo, ya que ellos mantienen la atención constante en la comunicación con el personal responsable de llevar a cabo las acciones para la

realización de las actividades en el desarrollo de proyectos.

4. Comprensión de la medida y la acción correctiva requerida por todo el personal  
Un KPI debe ser claro y preciso para el personal abarcando desde el objetivo con el que fue creado hasta las medidas que deben llevarse a cabo para corregir los posibles faltantes y con ello mejorar los estándares.
5. Atar la responsabilidad al individuo o equipo Un KPI es lo suficientemente profundo en la organización como para vincularlo con un individuo. En otras palabras, el CEO puede llamar a alguien y preguntar "por qué". Es decir, dar con la persona responsable de las acciones que no son 100 % satisfactorias, para que ella pueda ser orientada a la mejora.
6. significativo Cuando la gerencia del CEO y el personal se enfocan en el KPI, la organización obtiene los puntajes en todas las direcciones, por ejemplo: afecta la mayor parte de los factores críticos del éxito.
7. Impacto positivo El propósito para el que fueron creados los indicadores se da cuando existe una mejora en una medida clave por lo tanto existe un porcentaje de satisfacción del cliente y tendría un impacto positivo en muchas otras medidas: por ejemplo, afecta a todas las demás medidas de rendimiento en una manera positiva.

### 3.3. Pasos para crear KPIS

Los KPIS deben considerarse para el control y supervisión de las operaciones y deben enfocarse a los procesos críticos de la organización (Bishop, 2018) presenta pasos a considerar para la creación de indicadores clave de rendimiento.

- Número de KPI a formular.
- Frecuencia de medición de KPI.
- Si hay o no necesidad de KPI en todos los niveles organizacionales.

- Objetivos que se establecerán para los KPI identificados.
- Capacidad de gestión y controlabilidad de KPI.
- Diferenciar los KPI departamentales de toda la empresa.
- Infraestructura para soportar los KPI formulados.
- Identificar el propósito de establecer cada KPI.

Si estos puntos son exitosamente considerados se dará creación a indicadores efectivos. A continuación, se especifican de manera individual los puntos anteriores.

#### *Número de KPI a formular*

Una corporación puede enfrentarse a diversos problemas en su camino a la creación de indicadores, en algunas ocasiones recaen en el hecho de querer construir muchos indicadores para obtener suficiente información, lo cual puede conllevar a una sobrecarga de información que a su vez llevará a una extracción de información más compleja y por lo tanto el proceso de toma de decisiones se verá más complicado [1]. Por lo tanto, la forma más eficiente para establecer KPIS es en base a actividades críticas y es recomendable separar los KPI en partes o categorías tales como costo, calidad, entrega, seguridad, etc. (Barr, 2011) sugiere que en cada categoría el máximo de KPIS sea de tres.

#### *Frecuencia de medición de KPI*

La frecuencia de medición debe ser utilizada para determinar el nivel de rendimiento que llevan los indicadores y debe realizarse de manera continua, por lo tanto, es necesario establecer un periodo de medición: diario, mensual, trimestral o anual. Mediante esto se conocerá el desempeño obtenido durante el desarrollo de una actividad específica que pueda cumplir los objetivos establecidos. No obstante, los expertos en este tipo de mediciones dicen que “La verdadera mejora del proceso puede no ser visible durante semanas o meses.” y resaltan la importancia de “establecer un período y frecuencia para medir los indicadores clave de rendimiento que tengan sentido” [1]. (Kpilibrary, 2010) Menciona tres tipos de periodos de medición:

1. Períodos cerrados. Estos se llevan a cabo después de establecer tiempo determinado, es decir: la semana pasada, el mes pasado o el año pasado. Como productos finales se pueden obtener reportes o informes.
2. Períodos abiertos. Son utilizados para medir los objetivos del año a la fecha, pero carecen de estructura para el análisis de periodos anteriores.
3. Períodos en movimiento. Son los más orientados al futuro y son más conocidos como: los “Últimos 30 días, Últimos 90 días” un ejemplo es: “Número de cancelaciones en los últimos 30 días”.

*Si hay o no necesidad de KPI en todos los niveles organizacionales*

La definición de KPIS a nivel organizacional pueden contribuir al logro conjunto de los objetivos debido a que si existe una mejora por parte de los empleados esto determinara el buen rendimiento de la organización. Por lo tanto, es importante considerar establecer indicadores de rendimiento para el trabajo individual de los empleados y a su vez para el equipo, lo cual garantiza que el personal trabaje para el cumplimiento del plan estratégico en la organización.

*Objetivos que se establecerán para los KPI identificados*

La organización debe esclarecer los objetivos a los que desean llegar mediante la apropiada determinación de las áreas que requieren mejora. Existe un nivel de dificultad para poder medir los objetivos y es por ello por lo que se recomienda definir indicadores clave de rendimiento pequeños que sean fácil de cuantificar y a su vez estos deben ser actuales ya que es imposible que haya mejoras en el pasado.

*Diferenciar los KPI departamentales de toda la empresa*

En este enfoque se debe considerar establecer las diferencias entre los KPI institucionales y los departamentales, para las pequeñas organizaciones puede no ser necesario tener KPIS para todos los niveles, sin embargo, las medianas y grandes empresas deben clarificar el establecimiento de indicadores de rendimiento para todos los niveles de la organización, de lo contrario existirá una carencia de eficiencia. El beneficio obtenido de este punto es identificar con facilidad la responsabilidad de los

departamentos en caso de que se presente un incumplimiento de objetivos para poder resolverlos de la forma más adecuada y rápida posible.

*Infraestructura para soportar los KPI formulados*

El personal a cargo de la generación de KPIS debe considerar la relevancia que tiene el obtener la infraestructura necesaria para poder alimentar a los indicadores para obtener los beneficios adecuado y a su vez es necesario contar con fuentes de datos que sean capaces de proveer información suficiente y que sea la que reporta el personal a la organización para no sobrecargar a los empleados con actividades irrelevantes para los KPIS. De esta manera no existirá una ineficiencia en la infraestructura o en la información.

*Identificar el propósito de establecer cada KPI*

El nivel de capacidad para determinar el desempeño que tiene una organización es uno de los factores relevantes para identificar las áreas que requieren de la mejora y poder enfocarse en ellas para establecer adecuadamente los indicadores de rendimiento. Es por ello por lo que el personal al frente de la organización debe trabajar en conjunto para determinar el nivel de necesidad antes de la creación de los KPIS. (Thean, 2012) Establece cinco preguntas útiles para crear KIPS eficientes. La figura 3.8 ilustra los pasos a considerar para generar KPIS útiles.



**Figura 3.8:** Pasos útiles para la generación de KPIS.

### 3.4. Conclusión capítulo 3

Este capítulo es una pieza fundamental ya que como se planteó en el capítulo 1 la empresa cuenta con una certificación de calidad en CMMI y por lo tanto este análisis permite la identificación correcta de cada área de proceso con el fin de corroborar la importancia de algunos de los datos que puedan contribuir a la mejora de la toma de decisiones durante el desarrollo de proyectos, así como aprender a mejorar distintas actividades tomando en cuenta las lecciones aprendidas de los proyectos históricos. Por lo tanto, gracias a este análisis se pudo comprender la forma precisa de operar bajo los términos que se establecen, así como importancia de las actividades que se llevan a cabo para el cumplimiento de los objetivos que define la certificación y determinar qué datos son los más apropiados para poner en marcha el proceso ETL mediante la eficiencia en la creación de KPIS ya que algunos autores en sus investigaciones determinan las principales tareas para crearlas de la forma más apropiada.

## 4 Extracción Transformación y carga

En este capítulo se presenta el proceso ETL como tecnología para la creación eficiente de inteligencia de negocios, así como también los indicadores claves de rendimiento encontrados en la organización mediante la norma de calidad para el desarrollo de proyectos establecida de forma interna en la organización.

### 4.1. ETL en Business Intelligence

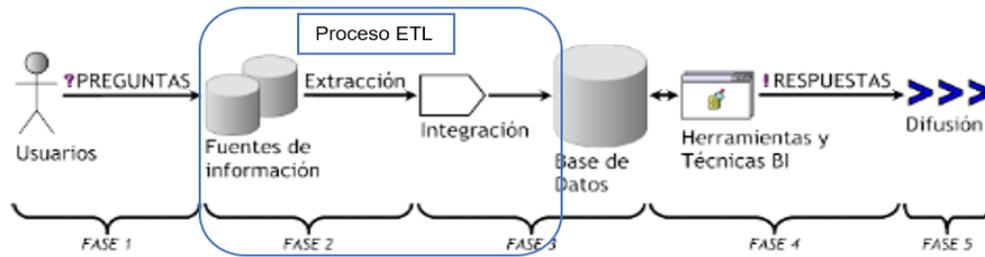
Dentro del contexto que involucra la creación de inteligencia de negocios existe todo un reto para el personal encargado de esta actividad debido a que esta tecnología debe ser orientada y personalizada en base a las necesidades de la organización, y es debido a ello que cada implementación es un mundo de cuestiones sin respuestas definitivas, simplemente se debe adaptar la tecnología a lo que sea más eficiente para la organización. El proceso de ETL es en si el primer paso para la creación eficiente de Inteligencia de negocios y sin duda alguna es una de las tecnologías más eficientes para incorporación de datos, es debido a ello que es la más utilizada para proporcionar información a los almacenes de datos y de esta manera alimentarlos.

Para obtener una correcta aplicación de inteligencia de negocios existen tres partes fundamentales que son:

- Proceso ETL: Extract Transform and Load por sus siglas en inglés; Es la primera etapa encargada del procesamiento e integración de las fuentes de datos
- Data warehouse: creación del almacenamiento de datos que va a recabar y guardar información después del proceso ETL.

- Explotación de información: en esta etapa final el objetivo es interactuar con herramientas de consulta y análisis que permitan generar informes finales.

Sin embargo (Ing. Bernabeu & Dario, 2007) representan mediante un diagrama las fases del proceso de aplicación de BI, esta imagen puede observarse en la figura 4.1 que se presenta a continuación, y enseguida se da una explicación de las fases que definen en su libro.



**Figura 4.1:** Fases del proceso BI.

Fase 1 – Dirigir y Planear. En esta fase los autores recomiendan el uso de la recolección de requerimientos para identificar de manera clara las necesidades que se deben abordar y posterior a ello se creen preguntas que sirvan de guía para el logro de objetivos.

Fase 2 – Recolección de Información. Mediante esta fase se debe realizar la extracción de datos de las diversas fuentes, datos o información que sirva para responder a las preguntas planteadas en la fase 1 debido a que son estas las que cubrirán los objetivos.

Fase 3 – Procesamiento de Datos. Aquí se lleva a cabo la integración y carga de los datos elegibles para poder ser analizados. Los autores dicen: “esta actividad puede realizarse mediante la creación de una nueva base de datos, agregando datos a una base de datos ya existente o bien consolidando la información”.

Fase 4 – Análisis y Producción. Una vez concluidas las etapas anteriores es necesario trabajar con la información que fue extraída para que sea integrada y continuar con el uso de herramientas diseñadas para inteligencia de negocios, en esta fase se generan reportes con los resultados de la tecnología.

Fase 5 – Difusión. En ella las partes interesadas obtendrán la información específica para comprender de forma coherente los datos finales que requieren de una forma

más eficaz.

## 4.2. Pasos para ETL

Para que el data warehouse pueda ser alimentado de manera correcta es necesario llevar a cabo la desnormalización e integración de la información, es decir agrupar los datos de las fuentes de datos de cada proceso de la organización, este proceso se define como: ETL extracción, transformación y carga de datos, es característico debido a la solución que se plantea para poder consolidar datos que puedan ser procesados de forma eficiente y rápida, está conformado por tres pasos principales; a continuación se presentan de forma separada y se mencionan algunas de las actividades importantes a considerar para obtener un mejor rendimiento en cada etapa.

### *Extracción*

Es el paso inicial por lo tanto es responsable de reunir la información que se tiene en las diferentes fuentes de datos, así como de recaudar datos diariamente por las organizaciones, generalmente son datos de texto plano como archivos de texto, hojas de cálculo, bases de datos transaccionales entre otros.



**Figura 4.2:** Datos importantes para el proceso de extracción.

La figura 4.2 Datos importantes para el proceso de extracción presenta tres tipos de datos que deben considerarse para obtener la mayor eficiencia del proceso:

1. Datos históricos: debido a que son estos los que pueden proporcionar información

para sustentar la evolución que tiene la organización por lo tanto es de suma relevancia considerar ese tipo de fuentes de información.

2. Datos actuales: es importante considerar información actual, datos (OLTP) que son los generados diariamente por las organizaciones.
3. Datos de diferentes formatos: para esta actividad es viable considerar todo tipo de información que puedan proporcionar todas las fuentes de almacenamiento existentes en la organización.

### *Transformación*

En un inicio después de detectar la información que se requiere se obtienen datos con diferentes formatos, debido a que estos se encuentran almacenados en diferentes formatos, bases de datos y hasta correos electrónicos, es debido a ello que se requiere de la limpieza absoluta de la información lo que llevará a la unificación del formato específico para los datos y así lograr la adaptación al procesamiento, de la misma manera debe considerarse que los datos sean congruentes así como compatibles con el almacén y es debido a ello que es necesario definir estándares a la información que deban ser integrados para normalizar de manera correcta. La figura 4.3 que se presenta a continuación representa algunas de las actividades más importantes de la transformación de datos.



**Figura 4.3:** Actividades de la transformación.

En la figura 5.2 se pueden observar cuatro tipos de actividades que se recomienda

utilizar para obtener la eficiencia adecuada de la actividad de transformación. A continuación, se presentan de forma enumerada para explicar su importancia.

1. Medida de atributos: para no exceder el límite de caracteres permitido se debe establecer el mínimo y máximo de atributos para los datos y con ello prevenir una sobrecarga.
2. Conversión de nombramientos: mediante la conversión se pretende establecer nombres a los datos que sean totalmente relacionados, pero sin exceder el límite de caracteres disponibles y que los nombramientos sean unificados en todo tipo de fuente.
3. Elegir la fuente más fiable: en algunos casos es necesario analizar la información ya que muchas veces es redundante debido a la cantidad de documentación proveniente de los archivos empleados para almacenamiento, en base a ello se debe elegir la fuente más viable para no caer en la reincidencia e incongruencia de datos.
4. Limpieza de datos: para culminar con el abastecimiento de información es necesario hacer una clasificación eficiente de los datos que aporten información valiosa y los que solo son complementarios para ello se requiere de la limpieza adecuada es decir excluir información incorrecta o inconsistente.

### *Carga*

El último paso es el aseguramiento de la integración de datos al almacén, para ello es necesario adecuar la estructura que tienen los datos para que sea la misma con la que fue diseñado el almacén, es por lo que se debe hacer una revisión previa al proceso total de carga. Debido a que los datos a almacenar representan información actual e histórica para la organización es una tarea con un periodo de tiempo extenso que puede consumir un esfuerzo representativo para el personal a cargo.

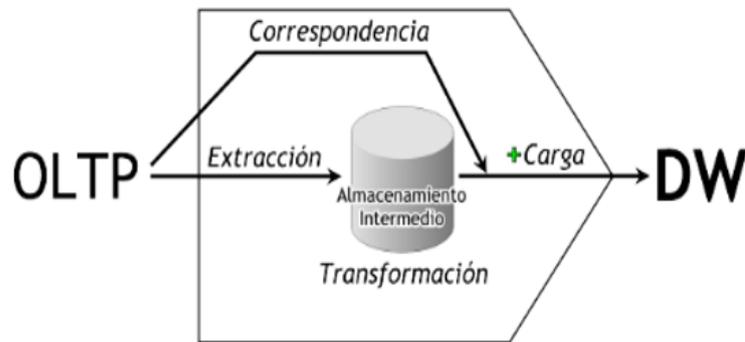
En la figura 4.4 se representan dos pasos importantes para el cumplimiento de la actividad:

1. Analizar los datos antes de la carga: mediante esta actividad se asegura que los datos sean los que aporten información adecuada, es decir los mejores datos.
2. Cargar la estructura del Data warehouse: concatenar la estructura de los datos con la del almacenamiento (data warehouse).



**Figura 4.4:** Actividades a considerar para la carga.

La siguiente figura 4.5. Interacción de los pasos para el proceso ETL fue definida por [8] y representa la interacción que existe entre las diferentes etapas que conforman el proceso.

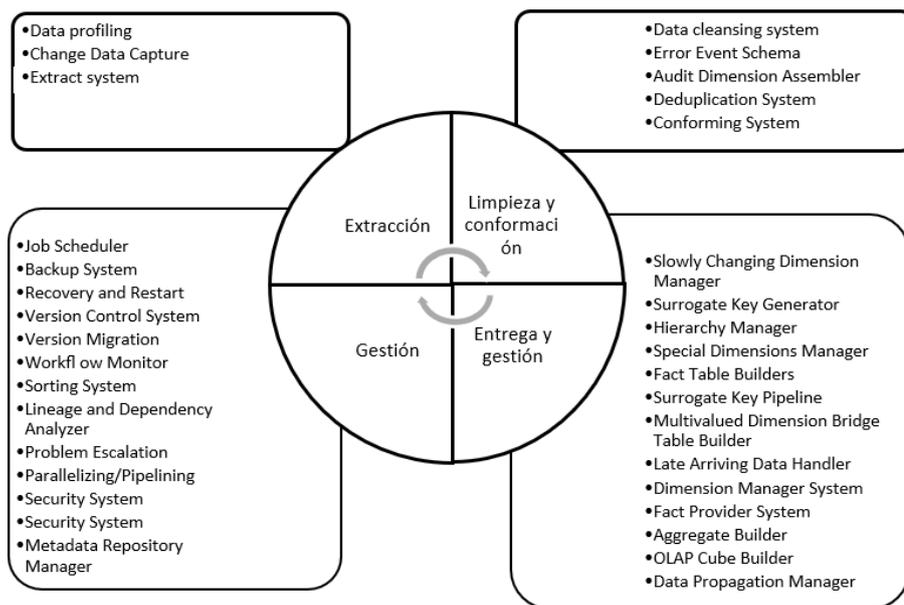


**Figura 4.5:** Interacción de los pasos para el proceso ETL.

En la imagen se puede observar como inicio las siglas Oltp que representan los distintos almacenamientos de datos que son comúnmente utilizados como fuentes de información, estas deben ser analizadas para la detección de información valiosa; posteriormente se encuentra el termino de extracción para dar a entender que

debe existir el proceso ya antes mencionado de análisis y limpieza de datos. Como paso siguiente la imagen ilustra un contenedor que representa a una base de datos intermedia encargada de recaudar toda esa información ya procesada pero creada con el propósito de no interferir con las actividades cotidianas de la organización, para finalmente ser la que establezca la comunicación directa con el data warehouse que va a fungir como almacenamiento principal para los datos ya procesados, eficientes que faciliten la toma de decisiones.

Sin embargo, diferentes autores han extendido los pasos en base a los estudios realizados para lograr una eficiencia completa del proceso, es por ello que (Conesa & Curto, 2015) hace una recopilación de actividades en su libro; esas actividades son definidas como subsistemas ETL y extraídas de Ralph Kimball que define treinta y cuatro clasificadores, que conforman cuatro grupos que se presentan a continuación en la figura 4.6. Los cuatro grupos del proceso ETL.



**Figura 4.6:** Los cuatro grupos del proceso ETL.

En la figura anterior se observan las cuatro actividades que define Ralph Kimball y los subsistemas por los cuales está conformado, a continuación, se muestra una

pequeña descripción de cada subsistema de acuerdo con el grupo al que corresponde.

### *Extracción*

- **Data profiling:** este primer subsistema pretende verificar la calidad de los datos mediante la indagación de ellos en las fuentes de almacenamiento para verificar que cumplan con los estándares definidos en un inicio como requerimientos.
- **Change Data Capture:** en este subsistema es necesaria la capturar los datos que han sufrido un cambio significativo pero no durante la carga general de todos los datos, sino desde la última actualización y únicamente cuando sean relevantes con ello se agilizará esta tarea sin saturar de información al almacén.
- **Extract system:** es todo un desafío lograr con éxito la extracción de datos, para permitir adquirir los datos de las fuentes de datos dispersas es necesario contar con herramientas para esta tarea, es por ello por lo que en casos complejos se requiere del desarrollo de algunas herramientas con lenguajes eficientes que personalicen el proceso para la organización.

### *Limpieza y Conformación*

- **Data cleansing system:** mediante ella es necesario poner en práctica la verificación de la calidad de datos y detectar las mejoras para el procesamiento de la información, así como omitir aquellos datos erróneos que puedan causar controversia en el proceso.
- **Error Event Schema:** el subsistema de rastreo de error pretende recaudar una relación de los inconvenientes que se presentan durante la transferencia de datos y a su vez alojara información para detectar en donde se produce, con ello mejorar la calidad y permitir la verificación de dichos inconvenientes.
- **Audit Dimension Assembler:** en el sistema ETL debe diseñarse un módulo de auditoría, el cual debe diseñarse para contener metadatos que ayuden a auditar las tablas de hechos y de esta manera conocer la calidad.

- **Deduplication System:** es difícil dentro del proceso contar con la capacidad para detectar información que se encuentra almacenada en diferentes partes para poder hacer una fusión de los datos y determinar qué información debe eliminarse y cual procesarse, es por ello que se sugiere el uso de patrones que ayuden a determinar las acciones que se deben llevar a cabo con respecto a la información duplicada.
- **Conforming System:** el sistema de conformación de información consiste en la conjunción de información mediante la detección de atributos en común para compartir datos entre tablas y poder integrar la información para mantener la estructura establecida en las tablas.

#### *Entrega y gestión*

- **Slowly Changing Dimension Manager:** dentro de este subsistema se generan la detección de cambios que ha sufrido la información y recae dentro de tres tipos de respuestas, la primera es la sobrescritura, la segunda agregar una nueva fila y la tercera agregar una columna; por lo tanto, este proceso es de suma importancia para la actualización de información.
- **Surrogate Key Generator:** el generador de clave sustituta debe asignar un número para que funja como clave principal en la fila de dimensión, independientemente de los activadores de base de datos.
- **Hierarchy Manager:** las jerarquías definen la dirección en la que los datos van a ser incrustados en base a la estructura establecida para la organización, por lo tanto, el sistema ETL mediante la utilización de las reglas debe completar la tabla de dimensiones correctamente.
- **Special Dimensions Manager:** la creación de un sistema de dimensiones especiales tiene la finalidad de reconocer las posiciones de la arquitectura ETL que sustente las características del diseño de la organización.

- Fact Table Builders: los constructores de tablas de hechos almacenan las mediciones de la organización. Las tablas de hechos se centran en la arquitectura ETL para que se puedan derivar los 3 tipos de tablas: hechos de transacción, hechos instantáneos periódicos e instantáneos acumulados.
- Surrogate Key Pipeline: el canal de claves permite la transformación de las claves asignadas en la operación a las claves aptas a la dimensión, es decir permite la correcta interpretación de las claves de asignación cuando ocurre un error en la integridad de la asignación, el ETL es el encargado de resolver este problema y cualquier otro aunado a la búsqueda o asignación de claves.
- Multivalued Dimension Bridge Table Builder: los constructores puente de las tablas de dimensión son uno de los principales objetivos que debe cubrir el proceso ETL debido a que es el encargado de mantener la comunicación eficiente entre tablas distintas.
- Late Arriving Data Handler: existe la posibilidad de que dentro del proceso ETL haya datos que toman mayor tiempo de procesamiento y es por ello por lo que llegan tarde, en estos casos es necesario hacer una modificación al proceso natural para procesar los datos rezagados por tardanza, por lo tanto es necesario hacer una búsqueda dentro del historial para identificar los datos que son funcionales cuando se realizó el procesamiento y con ello definir si deben ser tratados o verificados.
- Dimension Manager System: el administrador de dimensiones del sistema es el encargado de administrar las dimensiones mediante las actividades enumeradas a continuación: 1. Implementar etiquetas establecida en el diseño de la dimensión, 2. Agregar nuevas filas a la dimensión que puedan incluir nuevos datos, 3. Agregar nuevas filas para los cambios y mediante ello generar nuevas claves, 4. Modificar las filas para los cambios de los diferentes tipos sin cambiar las claves, 5. Actualizar el numero de la versión de modificación en las dimensiones, 6. Duplicar la dimensión aprobada a las tablas de hechos.

- **Fact Provider System:** proveedor de hechos es la persona responsable de la correcta administración de las tablas de hechos, así como su creación, actualización, mantenimiento y uso. Algunas de las tareas esenciales definidas por [19] son: 1. Recibir o descargar una dimensión replicada desde el administrador de aplicaciones; 2. Procesar los registros de dimensiones marcados como nuevos, pero con registros de dimensiones como nuevo, actualizar las claves, procesar los registros de las dimensiones marcados como nuevos, pero con fecha de caducidad; 3. Agregar las filas nuevas a las tablas de hechos; 4. Modificar las filas de las tablas de hechos para corregir errores y los cambios de dimensiones que llegan tarde. Entre otras.
- **Aggregate Builder:** los agregados dentro del proceso ETL son estructuras de datos específicas creadas para mejorar el rendimiento, con ello mantienen las filas de las tablas de hechos dentro de las tablas de dimensiones cuando lo necesitan las tablas de hechos agregadas.
- **OLAP Cube Builder:** los cubos OLAP son una parte externa del proceso ETL, estos cubos son alimentados de los esquemas relacionales y debido a que no proporcionan actividades como la limpieza son utilizados al final del proceso.
- **Data Propagation Manager:** para lograr la transmisión de los datos se requiere contar con el administrador de propagación de datos ya que es este el responsable de la presentación de los datos, integrados y conformes al almacén para que la información ya pueda ser manipulada por el personal involucrado y lograr los fines para los que fue destinada.

#### *Gestión*

- **Job Scheduler:** es importante para el proceso mantener un orden en las actividades a realizar por lo tanto debe ser administrado mediante el control de trabajo de acuerdo con la categoría a la que pertenece.
- **Backup System:** es de suma importancia dentro del proceso que exista la

posibilidad de poder obtener acceso a la información que fue o es modificada debido a que corre muchos riesgos que no siempre se pueden controlar y para ello se utilizan los respaldos del sistema que deben ser diseñados como parte del proceso.

- **Recovery and Restart:** para proteger el proceso de las fallas que puedan suscitarse es recomendable contar con un sistema de respaldo y recuperación del sistema que sea apto para la recuperación de fallos, detener y reiniciar y así minimizar el impacto.
- **Version Control System:** un complemento más que también es importante para el proceso es el control de versiones de los proyectos ETL y de los datos involucrados, es por ello que se debe archivar el control de los chequeos dentro y fuera para cada módulo para así permitir contrastar información.
- **Version Migration:** la migración de versiones permite la transición de proyectos en fase de desarrollo, hasta pruebas o hasta a producción, esta actividad necesita una interfaz que permita el control de versiones para administrar correctamente los cambios y en caso de ser necesario anular la migración. Esta operación debe tomar en cuenta todos los tipos de cambios presentados en los puntos anteriores, así como el uso de nuevas herramientas de BI y la evolución que puedan sufrir los informes.
- **Sorting System:** el sistema de clasificación es una de las capacidades fundamentales del proceso ETL, esta capacidad puede guiar al diagnóstico oportuno de inconvenientes en la calidad de datos y con ellos mejorar el rendimiento.
- **Lineage and Dependency Analyzer:** el analizador de linaje y dependencia es el que mostrará la evolución del procesamiento de los datos identificando las modificaciones en las que se encuentran envueltos los datos.
- **Problem Escalation:** el sistema de escalamiento de problemas es creado para asegurar la calidad antes de que los datos puedan ser manipulados por el

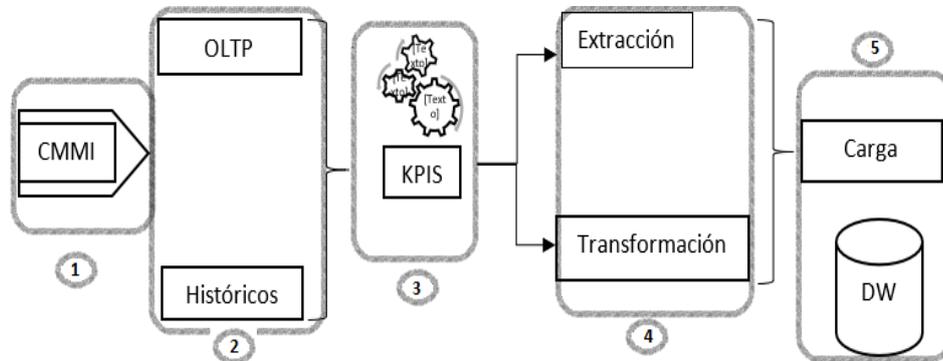
personal mediante las operaciones diarias, entonces si existe un error el proceso ETL debe hacérselo conocer al sistema de escalamiento de problemas para que haya una resolución mediante mensajes al personal encargado o al equipo que pueden ser desde correos electrónicos hasta mensajería a dispositivos personal en base a la importancia.

- **Parallelizing/Pipelining:** el surgimiento del sistema de paralelismo canalizado accede a que el sistema ETL haga entrega de la información en el rango de tiempo establecido es por ello por lo que aprovecha al máximo los recursos de computación paralela para cumplir en el tiempo necesario.
- **Security System:** el sistema de seguridad esta principalmente guiado al acceso del personal en el sistema ETL y al protocolo que se lleva a cabo para la transferencia de datos ya que no está por demás considerar un cifrado para poder realizar dicha actividad.
- **Compliance Manager:** el cumplimiento de almacén de datos sirve para la protección en la información ya que el proceso recoge datos confidenciales.
- **Metadata Repository Manager:** el administrador de metadatos debe implementar de las estrategias para el manejo de metadatos que fueron creados por el proceso y los que existían anteriormente.

### **4.3. Metodología ETL para la organización**

Como se presentó anteriormente existe una metodología establecida para la implementación de la tecnología BI y por lo tanto también la hay en la parte fundamental que es ETL, sin embargo esto no siempre es útil para cualquier organización es por ello que cada una debe adoptar las estrategias con mayor adecuación a su manera de operar, es debido a ello que para la resolución del problema planteado al inicio se llegó a la conclusión de modificar la estructura común del proceso y mediante

ello adaptar esta tecnología a la empresa. A continuación en la figura 4.7 se pueden observar los pasos definidos como metodología para la organización.



**Figura 4.7:** Proceso ETL para MBN

En la figura se presenta la estructura básica del ETL en la organización construida de 5 pasos básicos:

1. CMMI: como inicio se puede observar las siglas de la norma de calidad ya que es la encargada de la producción de información y la fuente principal para considerar.
2. OLTP e Históricos: OLTP representa a la información que se genera al instante así como las actualizaciones que se realizan a las fuentes de información para las actividades cotidianas del personal que labora en la organización, dentro de este mismo apartado se encuentra históricos que son aquellos datos que fueron almacenados con anterioridad y que pueden llevar a la mejora de la interpretación de los sucesos planteados con éxito o por lo contrario recaudar las lecciones aprendidas para no volver a recaer en los mismos errores, es por ello que se deben considerar en el proceso de toma de decisiones.
3. KPIS: son considerados dentro de la estructura debido a que son los encargados de detectar de manera oportuna la información necesaria que se requiere recaudar para la eficiencia del proceso de toma de decisiones.
4. Extracción y Transformación: la extracción que debe realizarse en base de

la información deseada evitando substraer información no deseada, es decir obtener datos precisos y claramente identificados que contribuyan a esclarecer las preguntas presentadas en el paso anterior mediante KPIS, la transformación debe llevarse a cabo en base a los lineamientos establecidos al inicio del proyecto para evitar la sobrecarga de información que haga a las actividades más pesadas y menos confiables.

5. Carga directa al DW: para finalizar la carga de datos, cabe destacar que originalmente en trabajos anteriores se maneja la creación de una base de datos intermedia encargada de almacenar los datos que fueron substraídos sin realizar un previo análisis, sin embargo para la solución de este proyecto se decidió omitir esta parte ya que mediante las KPIS se realizó un mejor análisis de la información y eso evita crear un doble almacenamiento por lo tanto la información pasaría directamente al data warehouse y con ello se omite la saturación de datos que implicaría la inversión de más tiempo en el análisis y con ello que cause mayor conflicto en el mantenimiento así como en el control eficaz de la duplicidad de información.

#### **4.4. Conclusión capítulo 4**

Este capítulo es dedicado en su totalidad a comprender aspectos fundamentales del proceso ETL como principal tecnología para el tratamiento de datos dentro de la inteligencia de negocios lo que significa que gracias a esto se pueden comprender las actividades esenciales que se deben llevar a cabo dentro del proceso así como identificar de manera clara y oportuna en donde se requiere de la intervención para adecuar el modelo y orientarlo a las actividades e información que se producen en la organización, por último se presenta la metodología final para la propuesta de solución y se explica cuáles son las actualizaciones que se realizaron de acuerdo con los modelos ya establecidos.

## 5 Resultados de KPIS y Pruebas ETL

En este último capítulo se presentan los resultados del análisis para KPIS así como el análisis independiente de los procesos que suministran reportes para proveer contenido elemental en la tarea, así como algunas de las actividades que se realizaron como parte del proceso de desarrollo ETL y una posible solución para la implementación, cabe destacar que es una simulación del proceso general, ya que como se presenta en la metodología principal de BI para la organización el modelado del data warehouse y ETL deben trabajar en conjunto para lograr la coincidencia en la estructura de ambos diseños tratamiento de datos y almacén.

### 5.1. Resultado de KPI definidas en la organización según CMMI

En las siguientes tablas presentadas a continuación, Tabla 4.1. Estado actual del proyecto, Tabla 4.2. Quejas generadas durante el proyecto, Tabla 4.3. Habilidades para conformación del equipo, Tabla 4.4. Uso del proyecto, Tabla 4.5. Estimación de Tecnología, Tabla 4.2. Calidad del proyecto, Tabla 4.7. Aceptación del usuario, Tabla 4.8. Efectividad en la ejecución, Tabla 4.9. Facilidad de utilización y Tabla 4.10. Tasa de cumplimiento de actividades. Presentan algunos puntos esenciales que se deben comprender para ejercer una buena orientación a la implementación de KPIS.

**Tabla 5.1:** Estado actual del proyecto.

¿Cuál es el estado actual de un proyecto?	
Objetivo Negocio	Identificar la fase en la que se encuentra un proyecto.
Necesidad de información	Para conocer el estado actual de proyecto y poder ejercer una buena toma de decisiones es fundamental conocer la fase de desarrollo en la que se encuentra el proyecto, ya que en base a esto se puede determinar un orden de prioridad para el proyecto.
Nombre Indicador	Estado de Fase de Proyecto.
Descripción	El estado del proyecto es considerado según el porcentaje de avance correspondiente a cada fase de proyecto.
Análisis	Cuando el indicador del tiempo real tiene un porcentaje mínimo comparado con el valor del indicador estimado, será necesario emitir una alerta para al responsable de proyecto ya que este es el encargado de realizar las acciones correspondientes.
Unidad de Medida	Porcentaje.
Destinatarios	Gestión de Negocios.
Fuentes	Datos planeados reales de esfuerzo del Plan Integral de Proyecto (Cronograma y Reporte de Actividades).
Procedimiento	Se asigna un porcentaje de avance por cada fase completada del ciclo de vida de proyecto, esto da un total de avance para el proyecto, al finaliza se obtendrá un 100 por ciento.
Periodicidad	Semanal.
Estado	Activo.
Responsable	Líder de Proyecto.
Frecuencia de medición	A la fecha.
Criterio de éxito	Rojo, Amarillo, Verde.

**Tabla 5.2:** Quejas generadas durante el proyecto.

¿Cuántas quejas se han generado en el transcurso del desarrollo del proyecto?	
Objetivo Negocio	Identificar qué tan conforme está el cliente con el desarrollo del proyecto.
Necesidad de información	Para conocer el estado actual de proceso y poder ejercer una buena toma de decisiones es fundamental conocer la fase de desarrollo en la que se encuentra el proyecto, ya que en base a esto se puede determinar una acción con prioridad para esclarecer la situación antes de que finalice el proyecto.
Nombre Indicador	Inconformidades del cliente.
Descripción	Controlar la intensidad en la generación de quejas para mejor la comunicación con el cliente.
Análisis	Cuando el indicador tiene un porcentaje elevado de acuerdos no satisfechos hacia el cliente es necesario identificar cuáles son las de mayor grado para que no recaiga en una mala experiencia con el cliente.
Unidad de Medida	Porcentaje.
Destinatarios	Gestión de Negocios y Gestión de Proyectos.
Fuentes	Plan Integral de Proyecto Reuniones de seguimiento con el cliente y con la gerencia.
Procedimiento	Se contabiliza la cantidad de inconformidades del cliente y se asigna un grado de relevancia, en caso de que la molestia se presente un foco rojo será necesario llegar a un acuerdo con el cliente para solucionar el problema.
Periodicidad	Semanal.
Estado	Activo.
Responsable	Líder de Proyecto.
Frecuencia de medición	Últimos 30 días.
Criterio de éxito	Rojo, Amarillo, Verde.

**Tabla 5.3:** Habilidades para conformación del equipo.

¿Cuáles son las habilidades que se requieren para el equipo de trabajo?	
Objetivo Negocio	Identificar al mejor prospecto para que pueda ser incorporado al equipo de trabajo en el desarrollo de proyecto.
Necesidad de información	Es importante conocer el perfil de la persona que va a ser incorporada al proyecto, así como sus habilidades para trabajar en equipo o con algún tipo de tecnología específico para determinar de qué forma aportara al desarrollo y logro de objetivos.
Nombre Indicador	Perfil personal.
Descripción	Controlar el acceso del personal por el tipo de proyecto a desarrollar.
Análisis	El indicador identificara un buen perfil de acuerdo con el porcentaje de habilidades de la persona.
Unidad de Medida	Porcentaje.
Destinatarios	Gestión de Negocios y Recursos Humanos.
Fuentes	Descripción del proyecto, plan de Desarrollo, plan general, recursos humanos.
Procedimiento	Se clasifica a la persona y se le asigna un porcentaje de acuerdo con las especificaciones del proyecto esto en base al nivel de manejo de lenguajes de programación, habilidades y actitudes, posteriormente se arroja un resultado de búsqueda para determinar a la persona más apta para ocupar el puesto.
Periodicidad	Cada inicio de proyecto.
Estado	Activo.
Responsable	Líder de Proyecto.
Frecuencia de medición	Histórico.
Criterio de éxito	Rojo, Amarillo, Verde.

**Tabla 5.4:** Uso del proyecto.

¿Cuál es la adecuación del proyecto al uso?	
Objetivo Negocio	Identificar el porcentaje de requisitos exitosos.
Necesidad de información	Para conocer el porcentaje de cumplimiento de los requisitos en el proyecto y que tan bueno es en cuestión del para que fue diseñado.
Nombre Indicador	Adecuación al uso.
Descripción	La adecuación al uso es determinada mediante el análisis de la funcionalidad del proyecto es decir si lo que realiza esta acorde a lo que fue diseñado.
Análisis	Es importante tener conocimiento de que tan certero es el diseño que hace el equipo para el apropiado diseño del software ya que se pueden presentar ocasiones en que el proyecto sea muy bueno, pero no haga lo que el cliente requiere, es debido a ello que se requiere determinar la adecuación del proyecto al uso.
Unidad de Medida	Numérico entero.
Destinatarios	Gestión de Negocios.
Fuentes	Especificación de requisitos, solicitud de cambios, auditoria de proyecto funcional sección de requerimientos, matriz de trazabilidad.
Procedimiento	Se determine un numero numérico en una escala del 1 al 10 en donde el 10 indica que es excelente la adecuación que se emplea en el uso del software.
Periodicidad	Semanal.
Estado	Activo.
Responsable	Líder de Proyecto.
Frecuencia de medición	Histórico.
Criterio de éxito	Rojo, Amarillo, Verde.
Fórmula	Adecuación = Requisitos cumplidos - Requisitos no cumplidos.

**Tabla 5.5:** Estimación de Tecnología.

¿Qué tipo de tecnología funcionó mejor para el desarrollo de un proyecto específico?	
Objetivo Negocio	identificar el mejor tipo de tecnología implementada para un proyecto anteriormente desarrollado y que puede colaborar a la mejora del proyecto.
Necesidad de información	Es importante identificar con que tecnología se ha tenido una buena experiencia de desarrollo, así como la que ocasione menos problemas de inestabilidad.
Nombre Indicador	Estimación Tecnología.
Descripción	Mediante la retroalimentación de la información se logrará controlar la tecnología de desarrollo en base a proyectos históricos y a las lecciones aprendidas durante la implementación.
Análisis	El indicador arrojará porcentajes de satisfacción en base al desarrollo y la satisfacción de cliente para continuar también con el mantenimiento.
Unidad de Medida	Porcentaje.
Destinatarios	Gestión de Negocios y Administración de Proyecto.
Fuentes	Lecciones aprendidas, Especificación de requerimientos, plan de pruebas del sistema.
Procedimiento	Se corrobora la información utilizada en la especificación de requerimientos acorde a las lecciones aprendidas y las pruebas del sistema. En base esto se asigna un porcentaje de viabilidad para los aspectos de mantenimiento, usabilidad y desarrollo.
Periodicidad	Cada inicio de proyecto.
Estado	Activo.
Responsable	Líder de Proyecto.
Frecuencia de medición	Histórico.
Criterio de éxito	Rojo, Amarillo, Verde.

**Tabla 5.6:** Calidad del proyecto.

¿Cuál es la calidad del proyecto concluido?	
Objetivo Negocio	Identificar la fase en la que se encuentra un proyecto para determinar la calidad del desarrollo en base al tiempo de cumplimiento de actividades.
Necesidad de información	Es importante determinar la calidad en la que se encuentra el proyecto, en determinado momento para cualquiera de los proyectos que se encuentren en marcha. La calidad es medible en base la tasa de cumplimientos en base al tiempo establecido de desarrollo.
Nombre Indicador	Estado de Fase de Proyecto.
Descripción	La calidad del proyecto es determinada mediante la fórmula presentada en el punto anterior es decir se realiza una comparación de actividades en base a lo planeado y real para posteriormente hacer una comparación más contra las fechas de entrega determinadas para el líder de proyecto, así como también al cliente.
Análisis	Cuando el indicador determina que existe un nivel escaso de calidad de entrega, será necesario emitir una alerta para al responsable de proyecto para que pueda intervenir en el proceso y poder guiar al cumplimiento eficiente de calidad de entrega.
Unidad de Medida	Numérico entero.
Destinatarios	Gestión de Negocios.
Fuentes	Seguimiento de hitos, alcance wbs, datos planeados y reales de esfuerzo del Plan Integral de Proyecto (Cronograma y Reporte de Actividades).
Procedimiento	Se asigna un numero al cumplimiento de actividades mediante una comparación de lo planeado contra lo real, para posteriormente hacer una comparación más contra las fechas de entrega determinadas para el líder de proyecto, así como también las del cliente y se asigna un numero terminado por el cumplimiento presentado posteriormente se requiere de una resta que pueda determinar un numero entero en el cual el mejor caso para calidad sea 1 y el peor sea 10 o más.
Periodicidad	Semanal.
Estado	Activo.
Responsable	Líder de Proyecto.
Medición	Histórico.
Criterio de éxito	Rojo, Amarillo, Verde.
Fórmula	Calidad= Tiempo y formalidad.

**Tabla 5.7:** Aceptación del usuario.

¿Cuál es el porcentaje de aceptación del usuario al proyecto?	
Objetivo Negocio	Identificar el porcentaje de requisitos exitosos.
Necesidad de información	Para conocer que tan contento está el usuario con su proyecto.
Nombre Indicador	Aceptación del usuario.
Descripción	La aceptación debe estar orientada a la aprobación del cliente.
Análisis	El análisis de la aceptación del usuario ayuda a determinar qué tan viable es el servicio de la empresa para satisfacer las necesidades del cliente.
Unidad de Medida	Numérico entero o decimal.
Destinatarios	Gestión de Negocios, APE, Fabrica de software.
Fuentes	Plan de proyecto y requerimientos.
Procedimiento	Una vez concluido el proyecto se debe llegar a una retroalimentación con el cliente que sea base para asignar un número de cumplimiento a los requisitos en una escala de 1 a 10 posteriormente se sumaran estas calificaciones para que se finalice con una división entre la cantidad de requisitos.
Periodicidad	Semanal.
Estado	Activo.
Responsable	Líder de Proyecto.
Medición	Histórico.
Criterio de éxito	Rojo, Amarillo, Verde.
Fórmula	$\text{Aceptación del usuario} = \frac{\text{Suma de Requisitos cumplidos}}{\text{número de requisitos}}$

**Tabla 5.8:** Efectividad en la ejecución.

¿Cuál es la efectividad en la ejecución de proyectos?	
Objetivo Negocio	Comprender de forma concisa la precisión de los acuerdos entre el cliente y el líder de proyecto para realizar los pagos establecidos durante el desarrollo del proyecto.
Necesidad de información	Para comprender el cumplimiento de los pagos en el proyecto solicitado y en caso de retraso tomar acciones preventivas que permitan llegar a un acuerdo.
Nombre Indicador	Efectividad del proyecto.
Descripción	La efectividad estas orientada al cumplimiento de los pagos durante el proyecto.
Análisis	La efectividad del proyecto debe estar orientada al cumplimiento en los acuerdos financieros con el cliente y debe involucrar acciones correctivas que permitan en determinado momento parar el desarrollo o continuar con acciones limitadas.
Unidad de Medida	Numérico entero o decimal.
Destinatarios	Gestión de Negocios, APE, Fabrica de software.
Fuentes	Acuerdos con el cliente.
Procedimiento	En el transcurso del desarrollo se requiere del conocimiento en el cumplimiento de los acuerdos financieros con el cliente por lo tanto se debe evaluar mediante una escala de 1 a 10 para determinar si existe un riesgo en la ejecución y si es así tomar las acciones que corresponden.
Periodicidad	Semanal.
Estado	Activo.
Responsable	Finanzas, Gestión de negocios y líder de proyecto.
Frecuencia de Medición	Mensual.
Criterio de éxito	Rojo, Amarillo, Verde.

**Tabla 5.9:** Efectividad en la ejecución.

¿Cuál es el porcentaje de facilidad de utilización de los proyectos?	
Objetivo Negocio	Comprender que tan efectivo es el análisis del equipo de trabajo para comprender los requerimientos del usuario y guiarlos a la óptima utilización.
Necesidad de información	Determinar si debe continuarse con las mismas prácticas para determinar el diseño del software.
Nombre Indicador	Facilidad de utilización.
Descripción	La utilización está orientada al conocimiento de la adaptabilidad del cliente al software.
Análisis	Para conocer la facilidad de utilización se requiere de la participación del cliente y es diagnosticada mediante las posibles causas por las que el usuario se sienta confundido en alguna acción específica con el software por lo tanto se debe determinar inicialmente por el diseño de los requerimientos durante el análisis de solución.
Unidad de Medida	Numérico entero o decimal.
Destinatarios	APE, Fabrica de software.
Fuentes	Encuestas de satisfacción del cliente.
Procedimiento	Mediante un análisis a la documentación del proyecto posterior a su culminación se debe determinar la claridad en el uso del software, en base a la determinación que debe obtenerse mediante la determinación y asignación de un numero en escala de 1 a 10.
Periodicidad	Cada revisión con el cliente.
Estado	Activo.
Responsable	APE, líder de proyecto.
Frecuencia de Medición	Histórico.
Criterio de éxito	Rojo, Amarillo, Verde.

**Tabla 5.10:** Tasa de cumplimiento de actividades.

¿Cuál es la Tasa de cumplimiento de actividades?	
Objetivo Negocio	Determinar el rendimiento del equipo con respecto a las actividades programadas al inicio del proyecto.
Necesidad de información	Determinar si se requiere tiempo extra para cubrir el desarrollo de actividades en el proyecto.
Nombre Indicador	Tasa de cumplimiento de actividades.
Descripción	Este indicador debe determinar si las actividades se están desarrollando dentro del tiempo establecido en el plan.
Análisis	Para determinar si el cumplimiento de actividades es efectivo se debe convocar al equipo de desarrollo del proyecto y determinar el avance personal con respecto a las tareas personales asignadas en un inicio.
Unidad de Medida	Numérico.
Destinatarios	Responsable de proyecto.
Fuentes	Actividades personales del equipo de desarrollo y plan integral de proyecto.
Procedimiento	La tasa de cumplimiento de actividades esta sujeta al avance que se tenga por parte del equipo, es por ello que se debe determinar mediante la documentación de las actividades del personal que integra el equipo de desarrollo y comparar el tiempo con respecto a la planeación inicial, si existen inconvenientes se debe determinar el nivel de alerta para este indicador y determinar las medidas adecuadas para solventar el incumplimiento de las actividades.
Periodicidad	Semanal.
Estado	Activo.
Responsable	Líder de proyecto.
Frecuencia de Medición	A la fecha.
Criterio de éxito	Rojo, Amarillo, Verde.

## 5.2. Áreas de proceso que proveen información para la solución de KPIS

Como parte del trabajo que comprende la definición de indicadores se realizó un análisis independiente a los seis procesos de CMMI para determinar la manera en que colaboran con cada KPI, la siguiente tabla 5.11 revela cuales son los procesos principales (destacados mediante una x) de los que se requiere información para apoyar al tratamiento de datos, ya que como se planteó en el capítulo anterior el análisis previo de la formación y por medio de KPIS es parte la propuesta de mejora y adaptación de inteligencia de negocios para MBN.

**Tabla 5.11:** Procesos que proveen información.

KPI	Procesos que proveen información					
	PP	PMC	REQM	PPQA	MA	CM
¿Cuál es el estado actual de un proyecto?	X	X	—	—	—	—
¿Cuántas quejas se han generado en el transcurso del desarrollo del proyecto?	—	X	X	X	—	—
¿Cuáles son las habilidades que se requieren para el equipo de trabajo?	X	—	—	—	—	—
¿Cuál es la adecuación del proyecto al uso?	X	—	—	X	—	—
¿Qué tipo de tecnología funcionó mejor para el desarrollo de un proyecto específico?	X	X	X	X	—	—
¿Cuál es la calidad del proyecto concluido?	—	—	—	X	—	X
¿Cuál es el porcentaje de aceptación del usuario al proyecto?	X	—	—	X	—	—
¿Cuál es la efectividad en la ejecución de proyectos?	X	X	—	X	—	—
¿Cuál es el porcentaje de facilidad de utilización de los proyectos?	X	—	—	—	—	X
¿Cuál es la Tasa de cumplimiento de actividades?	X	—	—	—	—	—

### 5.3. Extracción de KPIS con R

En esta sección se abordará el uso del Software de licencia libre R ya que mediante él se manipulan las fuentes de información para substraer datos que dan solución a las KPIS definidas. La figura 5.1 es una muestra del modelo de negocio que se realizo para cada indicador e ilustra los pasos que se deben seguir para llevar a la solución de la KPI de ejemplo: ¿Qué tan efectivo es un proyecto?; una de las inrrogantes más importantes dentro del caso de estudio principal que es el control de proyectos.

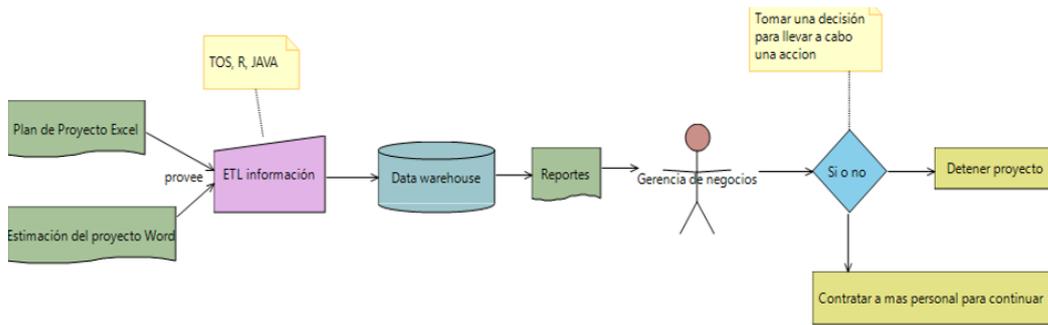


Figura 5.1: Diagrama KPI ¿Qué tan efectivo es un proyecto?

Mediante las Kpis presentadas con anterioridad se identificó que tipos de documentos y que información o campos son los más importantes para el procesamiento ETL, tomando en cuenta el estado natural de la información utilizada en la empresa. Las imagen que se presenta acontinuación es documentación que está almacenada directamente en los repositorios de la empresa por lo tanto es información original y contiene algunos errores que están fuera del alcance del equipo de ETL.

	A	B	C	D	E
1	Nombre	Iniciales	Rol	Abreviatura Rol	Responsabilidades
2	Patricia Trejo Xelhuantzi	PTX	Administrador del Proyecto Especifico	APE	Toma de decisiones, planeación
3	Gustavo Flores Sanchez	GFS	Responsable de Gerencia de Proyecto	RGPY	Mantener conocimiento y control de los
4	Alberto Portilla Flores	APF	Director de Operaciones	DO	Mantiene visibilidad sobre la asignación
5	Rosa Isaura Mejorada Sosa	RIMS	Responsable de Procesos y Calidad	CPC	Mantiene actualizados los procesos
6	Norma Sánchez Sánchez	NSS	Auditor	AU	Conocimiento de las diferentes fases del
7	Alberto Portilla Flores	APF	Medición y Análisis	MA	Conocer y valorar indicadores asociados
8	Norma Sánchez Sánchez	NSS	Administración de la Configuración	AC	Conocimiento y experiencia en el manejo de control de versiones y administración
9	Isaac Flores Hernández / Andrea Avila Palma	IFH / AAP	Responsable de pruebas	RP	Conocimiento y experiencia en la planeación y realización de pruebas de
10	Arnold Israel Muñoz Hernández	AIMH	Responsable de Desarrollo y Análisis	RDM	Conocimiento y experiencia en el
11	Erick Gabriel Morales Rodriguez	EGMR	Programador	AN	Conocimiento y experiencia en la obtención, especificación y análisis de
12	Arnold Israel Muñoz Hernández / Erick Gabriel Morales Rodriguez / Isaac Flores Hernández	AIMH / EGMR / IFH	Programador	PR	Conocimiento y/o experiencia en la programación, integración y pruebas unitarias.
13	Isaac Flores Hernández	IFH	Diseñador	DI	Conocimiento y experiencia en el diseño
14	Mara Gálvez Martínez / Andrea Avila Palma	MGM / AAP	Responsable de Manuales	RM	Conocimiento de las técnicas de redacción y experiencia en el desarrollo y
			Administrador de Base de Datos		Conocimiento en administración de

Figura 5.2: Documentación original MBN

### 5.3.1. Extracción de KPI 1

En las siguientes figuras 5.3 y 5.4 se muestra la información obtenida mediante la etapa de extracción para la primer KPI ¿Cuál es el estado actual de un proyecto? Para ello se requiere de la información específica encontrada en las secciones de Cronograma de actividades implícito en el documento del plan integral de proyecto y reporte de Actividades ambos de tipo xlsx. Los datos específicos son: Datos planeados y reales de esfuerzo.

```

# A tibble: 227 x 88
  "5. Cronograma" X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_10 X_11
<chr> <chr>
1 5.1 Supuestos de <NA> <NA>
2 "Las estimacione <NA> <NA>
3 <NA> <NA>
4 Fecha inicio de <NA> <NA> <NA> <NA> <NA> <NA> <NA> 41890 <NA> 41992 <NA>
5 No de Actividad <NA> Tipo- Desc- Resp- Recu- Hora- Hora- Fech- <NA> Fech- <NA>
6 <NA> <NA> <NA> <NA> <NA> <NA> <NA> Inic- Fin Inic- Fin
7 <NA> <NA> <NA> Plan- <NA> <NA> 20 19 <NA> <NA> <NA> <NA>
8 1 <NA> 18 Elab- PTX 1 3 2 41890 41890 41887 41887
9 2 <NA> 32 Vali- PTX,- 1 1 1 41890 41890 41890 41890
10 3 <NA> 13 Reun- PTX 1 4 3 41891 41891 41891 41892
# ... with 217 more rows, and 76 more variables: X_12 <chr>, X_13 <chr>,
# X_14 <dtm>, X_15 <dtm>, X_16 <dtm>, X_17 <dtm>, X_18 <chr>, X_19 <dtm>,
# X_20 <dtm>, X_21 <dtm>, X_22 <dtm>, X_23 <chr>, X_24 <dtm>, X_25 <dtm>,
# X_26 <dtm>, X_27 <dtm>, X_28 <chr>, X_29 <dtm>, X_30 <dtm>, X_31 <dtm>,
# X_32 <dtm>, X_33 <chr>, X_34 <dtm>, X_35 <dtm>, X_36 <dtm>, X_37 <dtm>,
# X_38 <chr>, X_39 <dtm>, X_40 <dtm>, X_41 <dtm>, X_42 <dtm>, X_43 <chr>,
# X_44 <dtm>, X_45 <dtm>, X_46 <dtm>, X_47 <dtm>, X_48 <chr>, X_49 <dtm>,
# X_50 <dtm>, X_51 <dtm>, X_52 <dtm>, X_53 <chr>, X_54 <dtm>, X_55 <dtm>,
# X_56 <dtm>, X_57 <dtm>, X_58 <chr>, X_59 <dtm>, X_60 <dtm>, X_61 <dtm>,
# X_62 <dtm>, X_63 <chr>, X_64 <dtm>, X_65 <dtm>, X_66 <dtm>, X_67 <dtm>,
# X_68 <chr>, X_69 <dtm>, X_70 <dtm>, X_71 <dtm>, X_72 <dtm>, X_73 <chr>,
# X_74 <dtm>, X_75 <dtm>, X_76 <dtm>, X_77 <dtm>, X_78 <chr>, X_79 <dtm>,
# X_80 <dtm>, X_81 <dtm>, X_82 <dtm>, X_83 <chr>, X_84 <dtm>, X_85 <dtm>,
# X_86 <dtm>, X_87 <dtm>

```

Figura 5.3: Datos Extraídos del cronograma de actividades.

```

# A tibble: 27 x 11
"Reporte de Activ- X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_10
<chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
1 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
2 Nombre Isaac <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
3 De Lunes: 41991 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
4 A Domingo: 41987 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
5 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
6 Proyecto Activi- Descrip- Lun Mar Mie Jue Vie Sab Dom Total
7 Sistema de Solici- 19.-Ju- Reunión- 1 <NA> <NA> <NA> <NA> <NA> <NA> 1
8 Sistema de Solici- 31.-Ve- Verific- <NA> <NA> <NA> 1.5 <NA> <NA> 1.5
9 Sistema de Solici- 33.-El- Reporte- <NA> <NA> <NA> 2.5 6 <NA> <NA> 8.5
10 Sistema de Solici- 31.-Ve- Verific- <NA> <NA> 2 <NA> <NA> <NA> 2
# ... with 17 more rows

```

Figura 5.4: Datos extraídos del reporte de actividades.

Como se puede apreciar existe un poco alteración en la información debido a que los datos encontrados con NA algunas veces son designados por el personal con relleno de color, o simplemente para designar que no aplica la información correspondiente.

5.3.2. Extracción de KPI 2

En la figura 5.5 que se presenta a continuación se puede observar el documento original que ayuda a responder la KPI ¿Cuántas quejas se han generado en el transcurso del desarrollo del proyecto?

Descripción	Responsable	Recursos	Asignados	Horas reales	Fecha Planeada		Fecha Real	
					Inicio	Fin	Inicio	Fin
<b>Planeación Inicial</b>					<b>20</b>	<b>19.00</b>		
Elaboración de plan general del proyecto	PTX	1	3	2.00	08/09/2014	08/09/2014	05/09/2014	05/09/2014
Validación/Aceptación del plan general del proyecto por parte del	PTX, RIMS	1	1	1.00	08/09/2014	08/09/2014	08/09/2014	08/09/2014
Reunión para reconocimiento de equipo de trabajo y distribución de	PTX	1	4	3.00	09/09/2014	09/09/2014	09/09/2014	10/09/2014
Definición de estructura del directorio del proyecto en el repositorio	PTX	1	2	3.00	09/09/2014	09/09/2014	09/09/2014	09/09/2014
Establecer permisos de acceso al equipo de trabajo a las carpetas	PTX	1	2	3.00	10/09/2014	10/09/2014	10/09/2014	10/09/2014
Entrevista inicial con el cliente para identificación de reglas de negocio	PTX	1	4	3.00	10/09/2014	10/09/2014	10/09/2014	10/09/2014
Realizar plan de desarrollo	PTX	1	1	1.00	10/09/2014	10/09/2014	10/09/2014	10/09/2014
Verificación del plan de desarrollo	PTX, HERH	2	1	1.50	10/09/2014	10/09/2014	10/09/2014	10/09/2014
Corrección de defectos encontrados en el plan de desarrollo	PTX	1	1	0.50	10/09/2014	10/09/2014	10/09/2014	10/09/2014
Validación/Aceptación del plan de desarrollo	PTX, RIMS	1	1	1.00	11/09/2014	11/09/2014	11/09/2014	11/09/2014
<b>Especificación de Requerimientos</b>					<b>163</b>	<b>156.00</b>		
Reunión para distribución de tareas de esta fase	PTX	1	1	1.00	09/09/2014	09/09/2014	09/09/2014	09/09/2014
Realizar levantamiento de requerimientos (En sitio con el cliente)	EGMR, IFH, RIMS	2	27.00	26.00	08/09/2014	11/09/2014	09/09/2014	18/09/2014
Realizar levantamiento de requerimientos (En sitio con el cliente)	EGMR, PTX	2	2.00	2.00	23/10/2014	23/10/2014	23/10/2014	23/10/2014
Realizar BPMN de las reglas de negocio	EGMR	1	6.00	4.00	10/09/2014	11/09/2014	10/09/2014	11/09/2014
Realizar lista de requerimientos	EGMR,	1	14.00	10.00	12/09/2014	18/09/2014	12/09/2014	18/09/2014
Realizar lista de requerimientos Solicitud de Cambios	EGMR,	1	0.50	0.50	24/10/2014	24/10/2014	24/10/2014	24/10/2014

Figura 5.5: Documentación del plan de proyecto.

200	MGM	1	2.250	2.00	41985
201	PTX, AAP	2	3.000	3.00	41982
202	PTX	1	1.000	1.00	41982
203	PTX	1	1.000	1.00	41985
204	PTX	1	1.000	0.50	41982
205	PTX, RIMS	2	1.500	2.00	41988
206	PTX, RIMS	2	1.500	2.00	41988
207	NSS	1	1.000	1.00	41982
208	<NA>	NA	55.000	48.25	<NA>
209	AAP, MGM	2	11.500	10.75	41983
210	IFH	1	1.250	1.00	41984
211	AAP	1	1.250	1.00	41984
212	IFH	1	8.000	5.00	41983
213	PTX, AIMH, MGM	3	1.000	1.50	41984
214	PTX, RIMS	2	1.000	1.50	41988
215	NSS	1	1.000	1.00	41984
216	PTX, RIMS	2	5.500	4.00	41988
217	PTX, RIMS	2	5.000	4.00	41988
218	EGMR, MGM, IFH	3	9.000	6.00	41981
219	AIMH, EGMR, PTX	3	2.500	3.50	41988
220	PTX	1	1.000	1.00	41982
221	PTX, AAP, MGM	2	6.000	7.00	41982
222	PTX	1	1.000	1.00	41989
X_1	Fecha Real	X_2			
1	Fin	Inicio	Fin		
2	<NA>	<NA>	<NA>		
3	41890	41887	41887		
4	41890	41890	41890		
5	41891	41891	41892		
6	41891	41891	41891		
7	41892	41892	41892		
8	41892	41892	41892		
9	41892	41892	41892		
10	41892	41892	41892		
11	41892	41892	41892		

Figura 5.6: Datos extraídos del plan de proyecto.

La figura 5.6 es una captura de los datos que se obtienen mediante la extracción de un documento excel para soporte del plan Integral de proyecto y acciones correctivas de tipo docx.

### 5.3.3. Extracción de KPI 3

Para la extracción de la información de la KPI 3 ¿Cuáles son las habilidades que se requieren para el equipo de trabajo? se realizó una conversión de formato de una extensión .DOCX a XML como prueba para extraer información desde r, los documentos que se deben procesar son: descripción del proyecto, plan de desarrollo y recursos humanos todos ellos con formato docx y el plan general con xlsx.

<p><b>3.1. Ciclos y Actividades.</b></p> <p><b>Inicio</b> Inicio de proyecto</p> <p><b>Especificación de Requerimientos</b> Levantamiento de requerimientos (En sitio con el cliente) Realizar especificación de requerimientos Verificación de especificación de requerimientos Corrección de especificación de requerimientos Validación/Aceptación de la especificación de requerimien Elaborar Matriz de Trazabilidad Realizar <b>baseline</b> de documentos de requerimientos Entrega de documento de especificación de requerimiento</p> <p><b>Análisis y diseño</b> Realizar diseño de arquitectura Realizar diseño detallado Realizar modelado de datos Realizar diseño de interfaz Verificación del documento de análisis y diseño Corrección de defectos encontrados al documento de aná Validación/Aceptación del documento de análisis y diseño Realizar <b>baseline</b> de fase de análisis y diseño Entrega de documento de análisis y diseño</p> <p><b>Construcción</b> Programar módulo de <b>Inicio de sesión</b> Programar módulo de <b>Viáticos</b> Programar sub módulo de <b>Solicitud de Viático</b> Programar sub módulo de <b>Comprobación de Gas</b> Programar módulo de <b>Aprobación de Solicitud de Viáticos</b> Programar sub módulo de <b>Verificación de Solicituc</b> Programar sub módulo de <b>Verificación de Gastos</b> Programar módulo de <b>Validación de Viáticos Aprobados</b> Programar sub módulo de <b>Concentrado de Solicit</b></p>	<pre> &lt;w:fldChar w:fldCharType="begin"/&gt; &lt;/w:r&gt; &lt;w:r&gt; &lt;w:rPr&gt; &lt;w:rStyle w:val="Nmerodepgina"/&gt; &lt;/w:rPr&gt; &lt;w:instrText xml:space="preserve"&gt;NUMPAGES&lt;/w:instrText&gt; &lt;/w:r&gt; &lt;w:r&gt; &lt;w:rPr&gt; &lt;w:rStyle w:val="Nmerodepgina"/&gt; &lt;/w:rPr&gt; &lt;w:fldChar w:fldCharType="separate"/&gt; &lt;/w:r&gt; &lt;w:r&gt; &lt;w:rPr&gt; &lt;w:rStyle w:val="Nmerodepgina"/&gt; &lt;w:noProof/&gt; &lt;/w:rPr&gt; &lt;w:t&gt;4&lt;/w:t&gt; &lt;/w:r&gt; &lt;w:r&gt; &lt;w:rPr&gt; &lt;w:rStyle w:val="Nmerodepgina"/&gt; &lt;/w:rPr&gt; &lt;w:fldChar w:fldCharType="end"/&gt; &lt;/w:r&gt; &lt;w:r&gt; &lt;w:rPr&gt; &lt;w:rStyle w:val="Nmerodepgina"/&gt; &lt;/w:rPr&gt; &lt;w:fldChar w:fldCharType="begin"/&gt; &lt;/w:r&gt; &lt;w:r&gt; &lt;w:rPr&gt; &lt;w:rStyle w:val="Nmerodepgina"/&gt; &lt;/w:rPr&gt; &lt;w:instrText xml:space="preserve"&gt;PAGE&lt;/w:instrText&gt; </pre>
---	---

Figura 5.7: Documentación del plan general y datos obtenidos.

La figura 5.7 que se pueden observar arriba son una representación de la información antes y después de la extracción en ellas se puede apreciar que existe una pérdida significativa de la información debido a la conversión, por lo tanto, no es recomendable realizar este procedimiento para toda la documentación.

5.3.4. Extracción de KPI 4

Para la KPI 4 ¿Cuál es la adecuación del proyecto al uso? Se presenta la siguiente figura 5.8 de la documentación original que fue analizada, los archivos que deben considerarse son: Especificación de requerimientos y solicitud de cambios ambos docx; auditoria de proyecto funcional (hoja de requerimientos) y por último matriz de trazabilidad ambos documentos con formato xlsx.

Figura 5.8: Documentación matriz de trazabilidad.

```
# A tibble: 21 x 9
  Cliente      `Rosa Isaura Mej~ X_1 X_2 X_3 X_4 X_5 X_6 X_7
<chr>         <chr>         <chr> <chr> <chr> <chr> <chr> <chr> <chr>
1 Proyecto    Sistema de Solicit~ <NA> <NA> <NA> <NA> <NA> <NA> <NA>
2 <NA>        <NA>          <NA> <NA> <NA> <NA> <NA> <NA> <NA>
3 Matriz de traza~ <NA>         <NA> <NA> <NA> <NA> <NA> <NA> <NA>
4 Necesidad   Objetivo      Foli~ SSVE~ Elem~ Docu~ Docu~ Caso~ Camb~
5 El sistema debe~ Tener control so~ RF00~ CU01~ 1.In~ 1. I~ CP001 CPS0~ <NA>
6 El sistema debe~ Realizar Solicit~ RF00~ CU02~ 2. V~ 2. S~ CP00~ CPS0~ <NA>
7 El sistema dent~ Control sobre lo~ RF00~ CU02~ 2. V~ 4. C~ CP004 CPS0~ <NA>
8 El módulo de v~ Tener un mejor a~ RF00~ CU02~ 2.Vi~ 2. S~ <NA> <NA> <NA>
9 El sistema debe~ Verificar las so~ RF00~ CU07~ 3.Ve~ 6. V~ CP00~ CPS0~ <NA>
10 El sistema en e~ Tener un control~ RF00~ CU11~ 3.1.~ 7. G~ CP005 CPS0~ <NA>
# ... with 11 more rows
```

Figura 5.9: Datos obtenidos de la matriz de trazabilidad.

Mediante la figura 5.9 se puede apreciar que no existe una pérdida de información sin embargo es necesario resaltar que el archivo contiene tablas en las que la información es muy precisa y clara por lo tanto contribuye al correcto procesamiento.

### 5.3.5. Extracción de KPI 5

Para esta KPI ¿Qué tipo de tecnología funcionó mejor para el desarrollo de un proyecto específico? Existe información inconclusa ya que no hay una actividad en la organización que pueda obtener las lecciones aprendidas de los proyectos históricos, es debido a ello que se propone la adopción de esta actividad definida en CMMI ya que es importante incluir este tipo de información para lograr obtener mejoras de los datos históricos de la organización. La información eficiente para extraer es: Lecciones aprendidas, Especificación de requerimientos y plan de pruebas del sistema la figura 5.10 permite apreciar las características de este formato y la figura 5.11 es el resultado que se obtiene del tratamiento de la información mediante la paquetería de ETL en R.

**1. Introducción**

Con el fin de evaluar el Sistema de Solicitud de Viáticos se plantea realizar los siguientes tipos de prueba:

- Pruebas Funcionales
- Pruebas de Interfaz de usuario.
- Pruebas de configuración.
- Pruebas de inicio de sesión.
- Pruebas de control de acceso y seguridad.
- Pruebas de carga.

El objetivo principal es probar la usabilidad del sistema. En nuestro caso, es determinar si un usuario puede utilizar el sistema completando satisfactoriamente el proceso que se desee.

**2. Descripción de pruebas**

Requisito	Tipo de requisito	Tipo de Prueba	Descripción de la Prueba	Resultado Esperado	Calendarización
CU01_Iniciar Sesión	Funcional	Prueba Funcional	Se debe ingresar el nombre de usuario y contraseña	Acceder al sistema y mostrar contenidos conforme a usuario.	11/12/2014
CU02_Viáticos	Funcional	Prueba Funcional.	Se deberá ingresar al módulo de Viáticos	Acceder al módulo con éxito.	11/12/2014

Figura 5.10: Documentación plan de pruebas.

```
[43] "PLAN DE PRUEBAS DEL SISTEMA"
[44] " Introducción"
[45] "Con el fin de evaluar el Sistema de Solicitud de Viáticos se
[46] "Pruebas Funcionales"
[47] "Pruebas de Interfaz de usuario."
[48] "Pruebas de configuración."
[49] "Pruebas de inicio de sesión."
[50] "Pruebas de control de acceso y seguridad."
[51] "Pruebas de carga."
[52] ""
[53] "El objetivo principal es probar la usabilidad del sistema.
[54] ""
[55] "Descripción de pruebas"
```

Figura 5.11: Datos obtenidos del plan de pruebas.

5.3.6. Extracción de KPI 6

El análisis de la KPI 6 ¿Cuál es la calidad del proyecto concluido? Determino que se llevara a cabo la solución mediante el procesamiento de los documentos: Seguimiento de hitos, alcance wbs, datos planeados y reales de esfuerzo todos estos integrados en el Plan Integral de Proyecto, específicamente en la hoja de Cronograma y Reporte de Actividades. La figura 5.12 que se presenta a continuación es una representación de la documentación original.

Etapa/Fase	Hitos/Entregables	Fecha Plan	Fecha Real	Estado	Observaciones
Especificación de requerimientos de software	Especificación de requisitos	03/10/2014	24/10/2014	Finalizado	
Especificación de requerimientos de software	Especificación de requisitos	30/10/2014	30/10/2014	Finalizado	Se entregará nuevamente con los cambios solicitados
Planeación	Plan Integral de Proyecto	28/09/2014	28/10/2014	Finalizado	
Análisis y Diseño	Análisis y Diseño	30/10/2014	11/11/2014	Finalizado	
Construcción	Código Fuente	01/12/2014	12/12/2014	Finalizado	
Integración y Pruebas	Manual de operación	15/12/2014	15/12/2014	Finalizado	
Integración y Pruebas	Manual de usuario	15/12/2014	15/12/2014	Finalizado	
Cierre	Manual de mantenimiento	15/12/2014	15/12/2014	Finalizado	
Cierre	Capacitación al usuario final	15/12/2014	15/12/2014	Finalizado	

Figura 5.12: Documentación seguimiento de hitos.

```
# A tibble: 90 x 10
  `4. Alcance del proy~ X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9
  <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
1 "Desarrollar el Siste~ <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
2 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
3 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
4 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
5 <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
6 4.1 Distribución de e~ <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA> <NA>
7 Fecha de actualización 41890 <NA> <NA> <NA> <NA> 1086~ <NA> <NA> <NA>
8 Etapa/Fase Acti~ Suba~ Esfu~ Esfu~ Esfu~ Esfu~ Esfu~ Disp~ Resp~
9 Planeación Inicial <NA> 20 20 20 0 20 19 <NA> <NA>
10 <NA> Elab~ 0.02 4 4 <NA> 4 3 <NA> <NA>
```

Figura 5.13: Procesamiento wbs e hitos.

La actividad de extracción arroja como resultado la figura 5.13, como se puede observar los datos sufrieron algunas alteraciones ya que el documento tiene caracteres especiales por el formato sin embargo los datos necesarios para el procesamiento no sufrieron alteración.

### 5.3.7. Extracción de KPI 9

Para lograr satisfacer la información relacionada con la KPI ¿Cuál es el porcentaje de facilidad de utilización de los proyectos? Se sugiere la utilización de encuestas al cliente, para adoptar las estrategias que ellos sugieran con respecto al diseño de la solución final, así como ver que tan efectivo es el manual para la interpretación del usuario, esto es deducible mediante la documentación que se presenta en los archivos de carta de aceptación que tiene formato docx y del cual se observa su contenido mediante la figura 5.14. y la figura 5.15 revela el resultado obtenido.

**1. Introducción**

Mediante este documento se acuerda con el cliente que el Sistema cumple con todas las funcionalidades requeridas y es de su entera satisfacción.

**2. Alcance**

El "Sistema de Solicitud de Viáticos" incluye los siguientes módulos:

1. Inicio de sesión
2. Viáticos (web y móvil)
  - a) Solicitud de Viático
  - b) Comprobación de Gastos
3. Verificación
  - a) Verificación de Solicitudes
  - b) Verificación de Gastos
4. Aprobación de Solicitud de Viáticos
  - a) Concentrado de Solicitudes
  - b) Realización de Pagos
5. Validación de Viáticos (web y móvil)
6. Configuraciones
  - a) Tabuladores de viáticos
  - b) Tabuladores de tipo de pagos
  - c) Reportes de gastos.

Figura 5.14: Documentación carta de aceptación.

```
[16] "Responsable:"
[17] "Patricia Trejo Xelhuantzi"
[18] "Cliente:"
[19] "Rosa Isaura Mejorada Sosa"
[20] ""
[21] ""
[22] ""
[23] "Copyright © 2014"
[24] ""
[25] ""
[26] "Contenido"
[27] ""
[28] " TOC \o \"1-3\" \h \z \u Documento de Aceptación PAGEREF_Toc265857373 \h 3"
[29] "1.Introducción PAGEREF_Toc265857374 \h 3"
[30] "2.Alcance PAGEREF_Toc265857375 \h 3"
[31] "3.Validación PAGEREF_Toc265857376 \h 4"
[32] "4.Soporte Técnico y Garantía PAGEREF_Toc265857377 \h 4"
[33] "5.Firma de Aceptación PAGEREF_Toc265857378 \h 4"
[34] ""
[35] ""
[36] ""
[37] ""
[38] "Documento de Aceptación"
[39] ""
[40] "Introducción"
[41] ""
[42] "Mediante este documento se acuerda con el cliente que el Sistema cumple con todas las funcionalidades requeridas y es de su entera satisfacción."
[43] ""
[44] "Alcance"
[45] "El "Sistema de Solicitud de Viáticos" incluye los siguientes módulos: "
[46] "1. Inicio de sesión"
[47] "2. Viáticos (web y móvil)"
[48] "Solicitud de Viático "
```

Figura 5.15: Datos obtenidos de la carta de aceptación.

5.3.8. Extracción de KPI 10

Se requiere del procesamiento de la documentación: Encuestas de satisfacción del cliente, reporte de seguimiento de todas las fases de desarrollo de software y por último, el reporte de actividades.xlsx que también es documentado por las fases del proyecto. Esta información dará solución al KPI 10 ¿Cuál es la Tasa de cumplimiento de actividades?, en la figura 5.16 se pueden observar el estado original de la información

REGISTRO DE ACTIVIDADES							
Actividad del Plan de Desarrollo	Fecha Inicio		Fecha fin		Esfuerzo en horas		Responsables
	Planeada	Real	Planeada	Real	Planeada	Real	
Reunión para la distribución de tareas	02/10/14	02/10/14	02/10/14	02/10/2014	1 hr	1 hr	Patricia Trejo Xelthuantzi APE
Reunión para la distribución de tareas	27/10/14	27/10/14	27/10/14	27/10/2014	1 hr	1 hr	Patricia Trejo Xelthuantzi APE
Realizar diseño de arquitectura	02/10/14	06/10/14	02/10/14	06/10/2014	25 hrs.	24 hrs	Erick Gabriel Morales Rodriguez AN, Arnold Israel Muñoz Hernández RDM
Verificación del documento de arquitectura	07/10/14	07/10/14	07/10/14	07/10/2014	2 hrs.	1hr 30 min	Mara Gálvez Martínez RE, Andrea Avila Palma RE
Corrección de defectos encontrados al documento de arquitectura	07/10/14	07/10/14	07/10/14	07/10/2014	2 hrs.	1 hr 30 min	Erick Gabriel Morales Rodriguez AN

Figura 5.16: Documentación Reporte de actividades.

La figura 5.17 se aprecia que existen datos faltantes, que son necesarios para la resolución, por lo tanto, no será de gran ayuda la asignación de datos aleatorios debido a que se requiere de información real para la eficaz toma de decisiones.

```
[69] ""
[70] ""
[71] "REGISTRO DE ACTIVIDADES"
[72] "Actividad del Plan de Desarrollo"
[73] "Fecha Inicio"
[74] "Fecha fin"
[75] "Esfuerzo en horas"
[76] "Responsables"
[77] ""
[78] "Planeada"
[79] "Real"
[80] "Planeada"
[81] "Real"
[82] "Planeada"
[83] "Real"
[84] ""
[85] "Reunión para la distribución de tareas "
[86] ""
[87] "02/10/2014"
[88] "02/10/2014"
[89] "02/10/2014"
[90] "02/10/2014"
[91] "1 hr"
[92] "1 hr"
[93] "Patricia Trejo Xelthuantzi APE "
[94] "Reunión para la distribución de tareas "
```

Figura 5.17: Datos obtenidos de seguimiento de actividades.

## 5.4. Conclusión capítulo 4

Gracias a este capítulo se puede determinar el logro alcanzado del proyecto ETL, así como los resultados de KPIS y el objetivo que tuvo su diseño para contribuir al análisis previo antes del procesamiento de los reportes que desprende CMMI y que tiene como fin la implementación de inteligencia de negocios en la empresa MBN.

Dentro del análisis que se realizó para la definición de KPIS se identificó de manera oportuna la información requerida que guía apropiadamente a la efectiva toma de decisiones este es el primer y más relevante análisis ya que de él pueden derivarse las consultas que se generan para ofrecer información determinada al usuario final.

Por lo tanto la propuesta de análisis que se presento en esta tesis es un primer paso, que incluso se tiene que llevar a cabo antes del procesamiento de la información que contienen los informes , para continuar con la extracción de la información “ R ” no permite realizar una extracción en base a este análisis ya que solo extrae los datos tal y como están en los documentos sin permitir la manipulación o selección previa de la información antes de almacenarlos en el Data Warehouse, de tal manera que si se desea continuar con la manipulación de este software se tendría que continuar con la metodología original del proceso ETL.

# 6 Conclusiones y trabajos futuros

## 6.1. Conclusiones

Para culminar con el trabajo de la realización de esta tesis se obtuvo un aporte significativo a la investigación del proceso ETL ya que hasta la fecha no se ha registrado una metodología similar a la propuesta en este trabajo, es decir que el proceso ETL se realice en base a la información que requieren los KPIS .

Con respecto al trabajo de ETL involucrado en alguna certificación de calidad tampoco existe algún registro, debido a que los aportes de ETL, KPIS y certificaciones de calidad son aportaciones totalmente independientes. Por lo tanto las certificaciones y la tecnología ETL solo se enlazan por medio de la generación de la información, lo cual conlleva a contribuir al proceso de ETL pero también implica un gran problema por la amplia generación de los informes que se establecen en este tipo de certificaciones.

En consecuencia fue todo un reto el poder manipular datos mediante la herramienta de tal forma que existe una complejidad incalculable en el tratamiento de los archivos que almacenan la información y es por ello que a lo largo de la tesis se aborda la relevancia e importancia que representan los reportes ya que son la principal y única fuente de información, sin embargo, el manejo de Rrepresenta una opción muy interesante para el procesamiento adecuado debido a la diversidad de los formatos que se manejan mediante la norma, desde contenido hasta estructura y a su vez son los que complican el procesamiento adecuado de la información, ya que al no contar con la misma estructura es muy complejo formular una misma solución ETL para todos esos documentos.

Cabe recalcar que mediante esta opción de solución que ofrece R es necesario el trabajo exhaustivo que implica el procesamiento manual de cada documento lo cual llevaría a un proceso muy lento y complejo a causa de que en la compañía se manejan de entre 700 a 2000 archivos por cada proyecto en desarrollo sin tomar en cuenta que esto es solo para el control de proyectos en base a la norma CMMI y que evolucionan en el transcurso de las actividades de creación y sin involucrar a MoProSoft la otra certificación utilizada en la empresa.

Para poder llevar a cabo una mejor solución se propone la unificación de archivos y el uso de una herramienta ETL presentadas en el capítulo 2 de esta tesis o en dado caso continuar con la adopción del software R, aunque, si no existiera la igualdad de estructuras en la documentación puede conllevar a recaer en este mismo problema.

### **6.2. Trabajos Futuros**

El gran problema que implica la creación personal de un sistema de procesamiento ETL para una organización que maneja dos tipos de certificaciones deja como aprendizaje las posibles investigaciones y aplicaciones del proceso ETL es por ello que en esta sección se pretende dar una pequeña aportación de las posibles mejoras que se pueden implementar a este proyecto:

1. Como primer punto y esencial es necesario mencionar que aún queda una brecha abierta a la oportunidad de la creación de un algoritmo que ayude mediante minería de datos o alguna otra tecnología a detectar los datos relevantes que pueden contener los archivos de texto plano ya que en esta situación fueron a partir de los formatos variantes y cambiantes constantemente. Sin tener que diseñar un algoritmo para cada archivo, lo cual generara trabajo excesivo para el desarrollador o equipo de desarrollo.
2. Como segundo punto es necesario verificar la importancia de la creación de una

base de datos intermedia que sea un almacén entre el procesamiento ETL y el data warehouse ya que esto permitirá tener un respaldo de la información que fue anteriormente procesada, y que a su vez servirá para la verificación o modificación de datos antes de almacenarlos de forma definitiva en el almacén principal.

3. Como tercer punto existe una oportunidad de investigación y aplicación en tiempo en el que se pretende llevar a cabo el procesamiento de la información ya que se complica el hecho de realizar esta actividad en tiempo real ya que en este caso los documentos manejados por la empresa muchas veces son modificados durante la realización de las actividades cotidianas por los empleados para desarrollar su trabajo de manera exitosa, es por ello que se sugiere organizar la información para poder llevar al éxito a este proceso.

## 7 Apendice



**International Journal of Science and Research (IJSR)**  
[www.ijsr.net](http://www.ijsr.net)  
ISSN (Online): 2319-7064

**Notification of Acceptance**

Date: 06/04/2018

Dear Raquel Contreras, Edith Viquez , Alberto Portilla, Edmundo Bonilla,

Your article has been accepted for publication in International Journal of Science and Research (IJSR).

To confirm your publication in Volume 7 Issue 4, April 2018, We request you to complete Final Submission on or before 13/04/2018, 11 PM IST.

You may Order for Certificate Hard Copies at this link [Order Certificate Hard Copy](#)

If you have any queries or doubts or need Deadline Extension, then please feel free to contact us at [www.ijsr.net@gmail.com](mailto:www.ijsr.net@gmail.com)

---

**Important Information**

Paper ID : ART20181414  
Paper Title : Defining KPI'S for an SME According to CMMI Dev 2 and MoProSoft  
Payable Fee : \$ 145 USD + Transaction Charges (This Fee Excludes Hard Copy Cost)

**How to make Final Submission?**

1. Upload your Final Article and Payment Proof at this link only <https://www.ijsr.net/finalsubmit.php>

**How to Make Payment?**

[Click here to pay via Paypal](#)  
[Click here to know SWIFT details.](#)



# Defining KPI'S for an SME According to CMMI Dev 2 and MoProSoft

Raquel Contreras<sup>1</sup>, Edith Vazquez<sup>2</sup>, Alberto Portilla<sup>3</sup>, Edmundo Bonilla<sup>4</sup>

<sup>1</sup>Master in Computer Systems, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, México

<sup>2</sup>Master in Computer Systems, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, México

<sup>3</sup>Innovation and talent development center, Miracle Business Network, La Loma Xicoténcatl, Tlaxcala, México

<sup>4</sup>Department of Systems and Computing, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, México

**Abstract:** *A strategy with much growth in the technological sector is the application of business intelligence (BI) to discover relevant aspects inside of the organization, this article that is developed as a knowledge base obtained in an SME in the technology sector certified with the international standard CMMI and the Mexican standard MoProSoft, both destined for the assurance of quality and responsible for providing sufficient information to know the indicators of the projects developed, these indicators will give way to the formulation of KPI'S that will be used later in the stages of ETL, Dataware and Exploitation of information for the complete implementation of BI.*

**Keywords:** KPIS, BI, SME, CMMI, MOPROSOFT

## 1. Introduction

The Business Intelligence (BI) area arises from the needs of companies to plan their operations from past experiences and future projections; this may involve a decision-making process based on the information generated as part of the operations and the environment. BI proposes, as a strategy, the application of technology to streamline the handling of information and processing of the most relevant data of the company. The objective is to use computer tools together with strategies to discover useful information and business indicators that guide the decision-making process.

A "Key Performance Indicator" (KPI, key performance indicator) is a measure associated with the result of some key action that leads to the fulfillment of a certain objective. For example: What is the total number of customers satisfied with the product? This KPI can help measure the quality of service offered by the organization and know if there is a loss or gain of customers. Therefore, it can help detect threats in the non-compliance of the product development that generates financial losses and alerts the person to take the appropriate measures.

The KPIs are important to determine the necessary actions that lead to the success of a company given its compliance or non-compliance. To address this problem, the specific case of a SME (small and medium-sized company) dedicated to information technology (IT) sector was analyzed, in which it was decided to implement a BI strategy in order to strengthen its strategic planning and operation process. And To guarantee the success of each project developed in the company, we have two certifications both at a level 2 maturity: MoProSoft (Mexican standard for the IT sector) and CMMi Dev 2 (focused on the processes involved in the development of projects). Both certifications generate a quantity of relevant data and that is why we intend to make

use of KPIs in order to detect through these two standards important data that can contribute to the growth of the company.

## 2. Key Performance Indicators

KPIs measure the level of performance of a given process, focusing on "how" and indicating how effective the processes are, so that the target can be achieved. KPIs are financial or non-financial metrics, used to quantify objectives that reflect the performance of an organization, and which are generally included in your strategic plan. These are overturned in the Integral Scorecard that collects and displays them, usually with a color code (red, yellow or green), providing information on compliance and not on the fixed objective. KPI's are "communication vehicles"; that allow the management of the organization to communicate the mission and vision of the company to employees, directly involving all employees in achieving the strategic objectives of the company. (ISOTools, 2017)

For an online marketing manager who has to define KPIS for the company's website, some possible examples are presented (Gomez, 2015):

- Total number of unique visitors, analyze not only the volume but the type of traffic.
- Whether it is recurring traffic or not, it is important to know who and why you repeat a visit
- Average stay time, key factor to know if the generated traffic is of quality or not.
- Conversion percentage, if the objective is online sales or the capture of leads, we can establish KPI to measure the percentage of visitors who have "clicked".

Volume 7 Issue 4, April 2018

[www.ijsr.net](http://www.ijsr.net)

Licensed Under Creative Commons Attribution CC BY

Paper ID: ART20181414

DOI: 10.21275/ART20181414

488

### 3. Definition of KPI's for SME

#### 3.1 KPI'S for MoProSoft

The Process Model for the Software Industry (MoProsoft) in Mexico promotes the standardization of its operation through the incorporation of the best practices in software engineering and management. To determine KPI'S in MoProsoft you must first mention the categories that comprise it. In the company in which the analysis was carried out, there is a process area that does not establish the standard, "FINANCE", but this is essential for any decision making, according to the needs of the SME and based on the two certifications it has, to detect KPI'S in MoProsoft the following areas are taken into account.

- GN: Business Management.
- GP: Process Management.
- GR: Resource Management.
- GPY: Project Management.
- BSI: Goods, Services and Infrastructure.
- RHAT: Human Resources and Work Environment.
- CO: Knowledge of the Organization.
- GF: Financial Management.

Once the process areas to be worked under the MoProSoft standard have been determined, the templates dictated by the standard should be analyzed to determine which are the most important for decision-making. It is worth mentioning that what the norm establishes is not enough since each area manager carries out extra documents or activities to carry out the entire process. For example, the GPY KPI'S and Finance Management are presented below.

Considering the 4 business lines:

- Which of them generates more projects per year?
- What is the sales executive that generates the most opportunities for the SME?
- What are the most frequent customers of the SME?
- How much has been invested in travel expenses?
- How much was invested in payroll?
- How much is the profit of each of the projects developed in the last quarter?
- What was the money invested in the last 6 months?

These are just some examples of KPI'S for two specific areas. The values of the KPI'S are determined based on the information gathered in the SME, which has been generated by the experts in the areas. In addition, from time to time the area managers of the company take some time to determine the current situation of the SME, this consumes too much time, that is why the objective of these KPI'S is to make an adequate data extraction that favors the growth of the company and solves everything that concerns. Those responsible must be able to carry out the decision making.

#### 3.2 KPI'S for CMMI

The scheme of good practices for development (CMMI DEV) at its maturity level 2 establishes seven process areas through which the guidelines required for the efficient development of projects are established. Next, six of the

process areas levels 2 are listed because they are those that are established in the organization:

- PP-Project Planning
- PMC-Project Monitoring and Control
- REQM-Requirements Management
- PPQA-Process and Product Quality Assurance
- MA-Measurement and Analysis
- CM-Configuration Management

These processes bring together the tasks, objectives and goals of the processes involved in a project to analyze and document the aspects that may arise in the course of development, whether the events that affect or benefit the achievement of objectives contemplating the different stages.

Therefore, it is necessary to control projects to obtain relevant information to detect alerts in the evolution of development and for management it is vital to know what happens in the company to be able to exercise good decision making, delimit or maintain a project.

In order to analyze the life cycle of each project under development, KPI'S are required, that is, project indicators. These are obtained from the information generated in regard to the CMMI dev 2 standards, in order to obtain indicators, it is necessary to analyze each documented file in the process of developing a specific task.

The KPI'S of the DMS and APE area (areas defined under the MoProsoft standard) are identified under the CMMI process areas; therefore, the analysis of these 6 areas defines that for the effective control of projects it is important to respond to the following questions:

- What is the current status of a project?
- How many complaints have been generated during the development of the project?
- What are the skills required for the work team?
- How many people are needed to develop a project?
- What type of technology worked best for the development of a specific project?
- What is the effectiveness in the execution of projects?

Some of these questions require historical information generated in successful projects of the company, therefore, it is necessary to know the documentation of previously developed projects. In this way, it will fulfill the main objective that is the implementation of business intelligence oriented mainly to the resolution of the KPI'S presented in this article for both quality standards.

**4. Table of MoProSoft and CMMI relations for the SME**

information requested of each of the Processes were studied; this will allow the extraction of the specific information of each process.

For the design of this table, the relationship and the

**Table 1:** Proposal of KPI'S

KPI'S	CMMI					MOPROSOFT										
	PP	PMC	REQM	PPQA	MA	CM	GN	GP	GPY	GR	RHAT	BSI	CO	APE	DMS	FINANCES
1. What is the profit of each of the projects developed?	X		X						X						X	X
2. Who is the sales executive that generates more opportunities?	X								X		X				X	X
3. How much was invested in payroll in the projects?	X										X					X
4. What were the total revenues?																X
5. What is the total amount of the invoices charged?																X
6. How much is invested in Memberships?																X
7. How much is invested in training?																X
8. How much was invested in incentives?																X
9. How much is invested in IMSS payroll?																X
10. How much is estimated for gifts for outsourcing partners?																X
11. How much is estimated in gifts from the CAAS collaborators?																X
12. How much training was requested?											X					
13. How much training do the collaborators receive?											X					
14. What is the duration Period of the trainings for each of the areas?											X					
15. What is the average time to achieve training objectives?											X					
16. What is the turnover rate of personnel?											X					
17. What is the rate of absenteeism?											X					
18. What is the average time of covered vacancies?											X					
19. What is the cost and time for hiring?											X					
20. What is the Talent Retention Index?											X					
21. What is the variation of the programmed unit cost compared to the one executed?									X							X
23. How much external training has been taught?									X		X					
24. How much internal training has been imparted?									X		X					
25. What is the number of satisfied customers?		X	X	X					X					X		
26. What is the effectiveness in the execution of projects?	X	X		X										X	X	
27. What is the current status of a project?	X	X												X	X	
28. How many complaints have been generated during the development of the project?		X	X	X					X					X		
29. What are the skills that are required for the work team?	X										X			X		
30. What type of technology	X	X	X	X												



- [7] Gómez "¿Qué es un KPI en marketing?" URL: <https://laculturadelmarketing.com/que-es-un-kpi-en-marketing/> (2015)

#### Author Profile



**Raquel Contreras** obtained the title of Engineer in Information Technology and Communications from the Technological Institute of Apizaco (ITA), Apizaco, Tlax, Mexico, in 2017. She is currently studying the Masters in Computer Systems, at the Technological Institute of Apizaco.



**Edith Vazquez** obtained the title of Engineer in Information Technology and Communications from the Technological Institute of Apizaco (ITA), Apizaco, Tlax, Mexico, in 2017. She is currently studying the Masters in Computer Systems, at the Technological Institute of Apizaco.



**Dr. Alberto Portilla Flores** is professor at the Universidad Autónoma de Tlaxcala and Manager of the Innovation and Technology Development Center at Miracle Business Network. He had a posdoc position at the Laboratoire Franco-Mexicain d' Informatique et Automatique LAFMIA – CNRS UMI 3175. He obtained his Phd in 2010, Université de Grenoble (HADAS team, Laboratory LIG) and the Universidad de las Américas-Puebla (CENTIA), under the supervision of Prof. Dr. Collet, Dr. Vargas-Solar, Dr. Zechinelli-Martini and Dr. Garcia-Bañuelos. From 1999 to 2002, he worked at Kronon Inc. as engineering manager. In July 1999 he obtained his Masters Degree on Computer Science at the Universidad de las Américas-Puebla, under the direction of Dr. Ahuactzin. In 1995 he obtained his bachelor's degree in Computer Engineering at the Universidad Autónoma de Tlaxcala. His research is related to non functional behavior of information systems.



**Dr. Edmundo Bonilla Huerta (M'11)** He obtained the title of Lic. In Computing by the Technological Institute of Apizaco (ITA), Apizaco, Tlax., Mexico, in 1996. The degree of M. in C. in Computational Sciences by the ITA, Apizaco, Tlax., Mexico, in 1998. And the degree of Doctorate in Informatique by the Université d'Angers, France in 2008. His areas of research include: quantum computing, fuzzy logic type 2, artificial vision, data mining and machine learning

Volume 7 Issue 4, April 2018

[www.ijsr.net](http://www.ijsr.net)

Licensed Under Creative Commons Attribution CC BY

Paper ID: ART20181414

DOI: 10.21275/ART20181414

482



**CARTA DE ACEPTACIÓN**

Tlaxcala Tlax a 19 de mayo del 2017

**MTRO. FELIPE PASCUAL ROSARIO AGUIRRE**  
**DIRECTOR DEL INSTITUTO TECNOLÓGICO DE APIZACO**  
**PRESENTE**

Por medio del presente le comunico que el **C. Raquel Contreras De La Fuente** alumno del Instituto Tecnológico de Apizaco de la Maestría en Sistemas Computacionales, ha sido aceptado para realizar sus **estancias profesionales bajo el proyecto “UN MODELO DE ETL PARA EL MANEJO DE LAS GERENCIAS EN MBN”** en nuestra empresa **MIRACLE BUSINESS NETWORK SA DE CV.** en el periodo que comprende del 28 de marzo del 2017 al 28 de agosto del 2017.

El C. Raquel Contreras De La Fuente estará bajo la responsabilidad de la Mtra. Patricia Trejo Responsable de Conocimiento de la Organización de esta empresa, con el firme objetivo de poner en práctica sus conocimientos y desarrollar habilidades que engrandezcan su perfil profesional y faciliten su posicionamiento dentro del ámbito laboral.

Es para nosotros un placer trabajar con alumnos de excelencia, con deseos de superación y actitud positiva ante los retos globales que hacen de cada esfuerzo un paso más hacia el éxito.

ATENTAMENTE

**Lic. Ariana Ramírez Moreno**  
**Responsable de Recursos Humanos**  
**y Ambiente de Trabajo**



**Miracle Business Network, S.A de C.V.**  
Hacemos de la tecnología, su mejor aliado de negocio.

**19 MAY 2017**

R.F.C.MBN060316R19  
Calle 37 No. 216 B La Loma Xicohténcatl  
Tlaxcala, Tlax. C.P. 90070 Tel. 01 246 416 45 08



Asunto: CARTA DE LIBERACIÓN

Tlaxcala, Tlax., 11 de septiembre del 2017

**Mtro. Felipe Pascual Rosario Aguirre**  
DIRECTOR DEL INSTITUTO TECNOLÓGICO DE APIZACO  
Con atención a:  
**Dr. José Federico Casco Vázquez**  
JEFE DEL DEPARTAMENTO DE DIVISIÓN DE ESTUDIOS  
DE POSGRADO E INVESTIGACIÓN  
P R E S E N T E

Me permito comunicar a usted, que la **Ing. Raquel Contreras de la Fuente**, estudiante de la Maestría en sistemas computacionales en el Tecnológico Nacional de México, Instituto Tecnológico de Apizaco, con número de matrícula **M11370777**, ha concluido una **Estancia Técnica**, realizando el proyecto de **“CMMi-dev 2 como base del proceso ETL para la generación de conocimiento en la inteligencia de negocios”**., durante el período comprendido del 28 de Marzo del 2017 al 28 de Agosto del 2017, los días martes, miércoles y viernes con un horario de 9:00 a 18:00 horas, teniendo como jefe inmediato a la M. en C. Patricia Trejo Xelhuantzi.

Sin más por el momento, se extiende la presente carta de liberación para los efectos correspondientes.

ATENTAMENTE

  
\_\_\_\_\_  
Lic. Vidalia Flores Mojica

  
Miracle Business Network, S.A de C.V.  
Hacemos de la tecnología, su mejor aliado de negocio.

11 SEP 2017

R.F.C.MBN060316R19  
Calle 37 No. 216 B La Loma Xicohténcatl  
Tlaxcala, Tlax. C.P. 90070 Tel. 01 246 416 45 08

Responsable de Recursos Humanos y Ambiente de Trabajo.

Calle 37 # 216-B  
La Loma Xicohténcatl, Tlaxcala, C.P. 90062

[www.mbn.com.mx](http://www.mbn.com.mx)  
Tel: 01 (246) 41 64508

V11.0, CARTADELIBERACION-PRACTICAS/ESTADIAS-RHAT

Pág. 1 de 1



Tlaxcala, Tlax., 08 Marzo de 2018

**Asunto: Constancia de Satisfacción.**

**Mtro. Felipe Pascual Rosario Aguirre**

DIRECTOR DEL INSTITUTO TECNOLÓGICO DE APIZACO

Con atención a:

**AT'N: Dr. José Federico Casco Vázquez**

JEFE DEL DEPARTAMENTO DE DIVISIÓN DE ESTUDIOS

DE POSGRADO E INVESTIGACIÓN

P R E S E N T E:

Sirva la presente para enviarle un cordial saludo y notificarle que posterior a la recepción del proyecto de tesis de la **Ing. Raquel Contreras de la Fuente**, alumna de la Maestría en Sistemas Computacionales, con número de matrícula **M11370777** de la institución que usted destacadamente dirige, se incluyó con el proyecto que lleva como título: **"CMMI-DEV 2 COMO BASE DEL PROCESO ETL PARA LA GENERACIÓN DE CONOCIMIENTO EN LA INTELIGENCIA DE NEGOCIOS"**.

Siendo este desarrollado bajo la dirección del Dr. Edmundo Bonilla Huerta catedrático de la citada maestría.

En virtud de que sea cubierto satisfactoriamente los objetivos establecidos del citado proyecto.

Tenemos a bien dar constancia de que dicho proyecto de tesis cubre y satisface las expectativas planteadas al inicio de este proyecto.

Sin más por el momento, se extiende la presente carta de satisfacción para los efectos correspondientes.

ATENTAMENTE

**Lic. Vidalia Flores Mojica**

**Responsable de Recursos Humanos y Ambiente de Trabajo.**

Calle 37 # 216-B  
La Loma Xicohténcatl, Tlaxcala, C.P. 90062

V11.0, CARTA DE SATISFACCIÓN-RHAT

  
**Miracle Business Network, S.A de C.V.**  
Hacemos de la tecnología, su mejor aliado de negocio.

**08 MAR 2018**

R.F.C.MBN060316RI9  
Calle 37 No. 216 B La Loma Xicohténcatl  
Tlaxcala, Tlax. C.P. 90070 Tel. 01 246 416 45 00

[www.mbn.com.mx](http://www.mbn.com.mx)  
Tel: 01 (246) 41 64508

Pág. 2 de 6

## 8 Bibliografía

- Adnan, Ilham, A. A., Usman, S.(2017). Análisis de rendimiento para Extraer, Transformar y Cargar (ETL) con Apache Hadoop en almacenamiento NASutilizando ISCSI. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Antolic Z.(2006). Ericsson Nikola Tesla d.d. Obtenido de Ericsson Nikola Tesla d.d: <https://courses.cs.ut.ee/MTAT.03.243/2014spring/uploads/Main/KPI.pdf>
- Bansal, S. K. (2014). Hacia un marco de Extracción Semántica-Transformación-Carga (ETL) para la integración de Big Data. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Barr, S. (2011). <http://www.staceybarr.com>. Obtenido de 74 How Many KPIs Do You Need?: <http://www.staceybarr.com/measure-up/74-how-many-kpis-do-you-need/>
- Bishop, D. A. (2018). Key Performance Indicators: Ideation to Creation. *eee Engineering Management Review*, 13-15.
- Bustamante Martínez, A., Amaru Galvis, E., Gómez Flórez, L. (2012). Técnicas de modelado de procesos de ETL: una revisión de alternativas y su aplicación en un proyecto de desarrollo de una solución de BI. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Chain, R. e. (2016). retos-operaciones-logistica. Obtenido de retos-operaciones-logistica: <http://retos-operaciones-logistica.eae.es/la-tecnologia-oltp-que-es-y-que-puede-hacer-por-ti/>
- Chen, G., Baoran An, A., Yan, L. (2016). Un nuevo sistema ETL paralelo basado en agentes para datos masivos. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Chu, Y.-C., Kuo,, W.-T., Cheng, Y.-R. (2017). Herramienta de respuesta de análisis de metadatos de supervivencia SMART. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- CMMI Institute. (2017). Obtenido de CMMI Institute: <http://cmmiinstitute.com/capability-maturity-model-integration>

- Conesa, J., Curto, J. (2015). ¿Como crear un data werehouse? En J. Conesa, J. Curto, ¿Como crear un data werehouse? (págs. 45-67). Barcelona: Editorial UOC.
- Darshan M., T., Amit , G., Y P , K., C.K, B. (2010). Exceso de velocidad en el procesamiento de ETL en almacenes de datos utilizando uniones de alto rendimiento para la captura de datos modificada (CDC). [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Guru. (2017). Obtenido de Guru:  
<http://www.guru99.com/ultimate-guide-etl-datawarehouse-testing.html>
- Hamed, I. (2015). Un enfoque basado en conocimiento para un proceso ETL de calidad. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Hamizah, M. R., Mohd, S. O., Lizawati, M. Y. (2015). Importancia de la integración de datos y ETL en el marco de Business Intelligence para la educación superior. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- IBM, (2015-2017). Obtenido de IBM:  
<https://www.ibm.com/ms-en/marketplace/datastage>
- Informatica. (2016). Obtenido de informatica:  
<https://www.informatica.com/mx/products/data-integration/powercenter.html>
- Bernabeu, I., Dario, R. (2007). Hefesto: Metodología propia para la construcción de un data warehouse. En R. D. Ing. Bernabeu, Data warehousing: investigación y sistematización de conceptos (págs. 6-26). Cordoba, Argentina: GFDL – GNU Free Documentation License.
- Javier, P., (2007). Softqanetwork. Obtenido de softqanetwork:  
<http://www.softqanetwork.com/moprosoft-modelo-de-procesos-de-software>
- Jianwen, S., Yan, T. (2017). Inteligencia de negocios revisada. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- klipfolio.(2017).klipfolio.Obtenido de klipfolio:  
<https://www.klipfolio.com/resources/articles/what-is-a-key-performance-indicator>
- Kpilibrary. (2010). <https://kpilibrary.com>. Obtenido de Measurement frequencies and periods: <https://kpilibrary.com/topics/measurement-frequencies-and-periods>
- Loreto, P., Fonseca, F., Morais, A., Peixoto, H., Abelha, A., Machado, J.(2017). Mejorando la atención de maternidad con Business Intelligence. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).

- Mary Beth Chrissis, M. K. (2009). CMMI Guía para la integración de procesos. En M. K. Mary Beth Chrissis, CMMI Guía para la integración de procesos (pág. 704). Pearson Educación 2009.
- Mind Tools. (2018). Mind Tools. Obtenido de Performance Management and KPIs: <https://www.mindtools.com/pages/article/newTMM87.htm>
- Olap.com. (2016). Obtenido de Olap.com: <http://olap.com/learn-bi-olap/olap-bi-definitions/business-intelligence/>
- Oracle.(2009).Obtenido de Oracle: <http://www.oracle.com/technetwork/middleware/data-integrator/overview/index.html>
- Parmenter, D. (2007). Key Performance Indicators Developing, Implementing, and using winning KPIs. Estados Unidos de America: John Wiley Sons Inc.
- Powerdata. (2015). Obtenido de powerdata: <http://www.powerdata.es/data-warehouse>
- Ralph, K., Ross, M. (2013). Capitulo 19 ETL Subsystems and Techniques. En R. Kimball, M. Ross, The data warehouse Toolkit (págs. 443-496). United States of America: John Wiley Sons, Inc.
- Spalevic, P., Emir, P., Ilic, S., Mladen, V., Stefan, P. (2016). Automatización del proceso ETL en el sistema aislado de base de datos a pequeña escala, [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- talend. (2017). Obtenido de talend: <https://www.talend.com/download/talend-open-studio/>
- Tang Jun; Cui Kai; Feng Yu; Tong Gang. (2009). La herramienta de investigación y aplicación de ETL en Business Intelligence Project. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Tang, J., Cui, K., Feng, Y., Tong, G. (2009). Herramienta de investigación y aplicación de ETL en un proyecto de inteligencia de negocios. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- Thean, P. (2012). Rhythm Systems. Obtenido de 5 Simple Steps to Create Useful KPIs: <http://www.rhythmsystems.com/blog/bid/122008/5-simple-steps-to-create-useful-kpi-s>
- Vilanueva, J., Chavez, Li, X. (2011). El proceso ETL basado en ontología para la creación de almacen de datos ontológicos. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).
- <https://es.slideshare.net/JorgeCarlos3/comparativa-herramientas-etl>
- <http://www.tutorialr.es/es/historiader.html>